



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number 99026**

**TO: Ashwin Mehta**  
**Location: CM1/9E07/9E12**  
**Art Unit: 1638**  
**Monday, July 28, 2003**

**Case Serial Number: 09/991458**

**From: Toby Port**  
**Location: Biotech-Chem Library**  
**CM1-6A04**  
**Phone: 308-3534**

**toby.port@uspto.gov**

### **Search Notes**

Dear Examiner Mehta,

Here are the results of your search.  
Please feel free to contact me if you have any questions.

Toby Port

**THIS PAGE BLANK (USPTO)**



# STIC SEARCH RESULTS FEEDBACK FORM

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor  
308-4258, CM1-1E01

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 – Circ. Desk



**THIS PAGE BLANK (USPTO)**



STIC-Biotech/ChemLib

99026

Fr m: Mehta, Ashwin  
Sent: Wednesday, July 16, 2003 5:50 PM  
To: STIC-Biotech/ChemLib  
Subject: seq search

RECEIVED

JUL 17 2003

STIC-BIOTECH/CHENLIB  
(STIC)

STIC,

Please search the commercial and interference databases for the following from 09/991,458:

- 1) the nucleotide sequence of SEQ ID NO:1
- 2) the amino acid sequence of SEQ ID NO: 2

My mail room is 9E12, office 9E07, art unit 1638.

Thank you,  
Ashwin

Ashwin Mehta  
United States Patent and Trademark Office  
Biotechnology Patent Examiner  
703-306-4540

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: July 25, 2003, 22:33:23 ; Search time 7263 Seconds

(without alignments)  
11460.006 Million cell updates/sec

Title: US-09-991-458-1

Perfect score: 2860  
Sequence: 1 ggcacgaggtactgtgttg.....ataaaaaaaaaaaaaaaaaa 2860Scoring table: IDENTITY\_NDC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

```
GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vtl:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2860	100.0	2860	3	AF013963	AF013963 Manduca s
2	2423.4	84.7	2861	3	AF006063	AF006063 Manduca s
3	446.6	15.6	4113	3	AY060329	AY060329 Drosophila
4	391.8	13.7	2473	6	AY191697	AY191697 Sequence
5	391.8	13.7	2497	3	AY121662	AY121662 Drosophila
6	383.4	13.4	2030	6	AY191701	AY191701 Sequence
7	382.8	13.4	2099	3	AY071318	AY071318 Drosophila
8	378.6	13.2	2097	3	AY058396	AY058396 Drosophila
9	378.6	13.2	2108	6	AX191703	AX191703 Sequence
10	366.8	12.8	1996	3	AF369383	AF369383 Aedes aeg
11	274.6	9.6	66809	2	AC015177	AC015177 Drosophila
12	274.6	9.6	166854	3	AC007330	AC007330 Drosophila
13	274.6	9.6	196337	3	AC005894	AC005894 Drosophila
14	274.6	9.6	242172	3	AE003832	AE003832 Drosophila
15	270.2	9.4	18567	2	AC018319	AC018319 Drosophila
16	270.2	9.4	165852	3	AC010022	AC010022 Drosophila
17	270.2	9.4	176929	3	AC091203	AC091203 Drosophila
18	270.2	9.4	286784	3	AE003352	AE003352 Drosophila
19	245.8	8.6	17325	2	AC018164	AC018164 Drosophila
20	245.8	8.6	184621	3	AC007440	AC007440 Drosophila
21	245.8	8.6	233148	3	AE003822	AE003822 Drosophila
22	231.4	8.1	37319	2	AC014400	AC014400 Drosophila
23	231.4	8.1	167902	3	AC104604	AC104604 Drosophila
24	231.4	8.1	302786	3	AE003434	AE003434 Drosophila
25	229.2	8.0	70610	2	AC017561	AC017561 Drosophila
26	229.2	8.0	153352	3	AC093103	AC093103 Drosophila
27	229.2	8.0	202495	3	AC092246	AC092246 Drosophila
28	229.2	8.0	260967	3	AE003645	AE003645 Drosophila
29	229.2	8.0	303043	3	DROSADH05	AE003411 Drosophila
30	227.2	7.9	1890	5	AF318177	AF318177 Drosophila
31	225.2	7.9	3375	10	AF411042	AF411042 Mus muscu
32	222	7.8	2817	6	AR048966	AR048966 Sequence
33	222	7.8	2817	6	AR080685	AR080685 Sequence
34	222	7.8	2817	10	RATGLY	L21672 Rat mRNA se
35	217.4	7.6	1894	9	AF286026	AF286026 Macaca mu
36	216	7.6	2255	4	BT099198	BT099198 Bos taurus
37	214.6	7.5	3393	10	AF109072	AF109072 Mus muscu
38	214.4	7.5	3188	4	BT09607	BT09607 B. taurus mR
39	213.4	7.5	3738	9	AK096607	AK096607 Homo sapi
40	213	7.4	1873	10	AF109391	AF109391 Mus muscu
41	213	7.4	1888	10	MMU238309	AJ238309 Mus muscu
42	212.6	7.4	1985	10	RATDPRSP	M80233 Rat cocaine
43	211.8	7.4	1911	9	S80071	S80071 hPROR-brain
44	211	7.4	3404	10	RATDOPER	M80570 Rat dopamin
45	209.6	7.3	2722	10	RATLPTHA	M88111 Rattus norv

## ALIGNMENTS

```
RESULT 1
AF013963      2860 bp  mRNA  linear  INV 30-OCT-2001
LOCUS
DEFINITION  Manduca sexta amino acid transporter/amino acid-gated channel for
ACCESSION  AF013963
VERSION    AF013963.2 GI:16356923
KEYWORDS
SOURCE
ORGANISM   Manduca sexta.
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
            Sphingioidea; Sphingidae; Sphinginae; Manduca.
REFERENCE  1 (bases 1 to 2860)
            Feldman,D.H., Harvey,W.R. and Stevens,B.R.
```

gene	1. .2860
CDS	/gene="CAATCH1" 101. .2002

Query Match	100.0%;	Score 2860;	DB 3;	Length 2860;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2860;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 GGCACGAGGTTACTTTGTTGGAGAACTGTTGGCGGTGTCGATCGATATTTCTGAACACA 60  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 1 GGCACGAGGTTACTTTGTTGGAGAACTGTTGGCGGTGTCGATCGATATTTCTGAACACA 60  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
QY 61 ATACACTAAACACATTCGACGTGATTTGTGGACAAATGAAATGACGGCCCAAGTAA 120  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 61 ATACACTAAACACATTCGACGTGATTTGTGGACAAATGAAATGACGGCCCAAGTAA 120  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
QY 121 CGGCGGTTTGGAGTCTGTCGAGCCCAAGATGAAACCAAAACGATCATCGCAATAAATTT 180  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 121 CGGCGGTTTGGAGTCTGTCGAGCCCAAGATGAAACCAAAACGATCATCGCAATAAATTT 180  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
QY 181 ACCTCCAGCAATTAATAAGCGCGCTCTGGAATATATGATGACACGGCATTTAGAGCGTGA 240

Db	181	ACCTCCAGCAATATATTAAGGCGCTCGATTAATATAGATGACCGGACTTAGAGCTGA	240
Qy	241	ACGGCAGAAAGTATGGTATGCTCCACCAACATTGAATTCTTGATGTCTGCATCGCTAC	300
Db	241	ACGCCAGAAAGTATGGTATGCTCCACCAACATTGAATTCTTGATGTCTGCATCGCTAC	300
Qy	301	ATCCGTGCTTTGGGTAAAGTATGTTGGCGGTTCCCTTATATGCGGTACAGAAATGAGGAGG	360
Db	301	ATCCGTGCTTTGGGTAAAGTATGTTGGCGGTTCCCTTATATGCGGTACAGAAATGAGGAGG	360
Qy	361	TGCTTTCCTGTGCTCCATAGCTCATGCTCTTTTACTTGTGCGGAAGCGTGTACTATT	420
Db	361	TGCTTTCCTGTGCTCCATAGCTCATGCTCTTTTACTTGTGCGGAAGCGTGTACTATT	420
Qy	421	AGAGTGTGCTCGGAAATTCAGTTCAACAACTCGTTAAAGTTTGGTCAATTTCCACC	480
Db	421	AGAGTGTGCTCGGAAATTCAGTTCAACAACTCGTTAAAGTTTGGTCAATTTCCACC	480
Qy	481	GGCCATGAAAGTACTGGATACGCTCAAGCTGCGGCTGGCGTGTACATCCTGTCTACTA	540
Db	481	GGCCATGAAAGTACTGGATACGCTCAAGCTGCGGCTGGCGTGTACATCCTGTCTACTA	540
Qy	541	CGGTGTGATCTGTGGTCTCTGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	600
Db	541	CGGTGTGATCTGTGGTCTCTGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	600
Qy	601	ATGGGCTATTTTGCACGCTGAGTGGGAGAACTGCGTACCTGATGATCCAACTTGTCTGC	660
Db	601	ATGGGCTATTTTGCACGCTGAGTGGGAGAACTGCGTACCTGATGATCCAACTTGTCTGC	660
Qy	661	ATGAGTCACAAACATACCAATGTGTACCAAGAGTGTCAACTCTATTTTGTAGAAGAGT	720
Db	661	ATGAGTCACAAACATACCAATGTGTACCAAGAGTGTCAACTCTATTTTGTAGAAGAGT	720
Qy	721	TCGCCAAAGAGGATGAGTGAATGGAAGGAGTCTCGGTGCCCCCATCTGCTACTTGTGT	780
Db	721	TCGCCAAAGAGGATGAGTGAATGGAAGGAGTCTCGGTGCCCCCATCTGCTACTTGTGT	780
Qy	781	GTGTCTATTCATGCGATGAGCTCATGGTGTTCGAGAGTGTGCGCCGAGAGAGTCAAGAGTTC	840
Db	781	GTGTCTATTCATGCGATGAGCTCATGGTGTTCGAGAGTGTGCGCCGAGAGAGTCAAGAGTTC	840
Qy	841	CGGCAAAAGGCGCTACTTCTCGGCGCTCTTCCCATACGTTGTGATGATCATTATTCAT	900
Db	841	CGGCAAAAGGCGCTACTTCTCGGCGCTCTTCCCATACGTTGTGATGATCATTATTCAT	900
Qy	901	CACCAACAATCATCTGCCGAGTCACTAGCAGGATCTGTTCTTCGACAGCCCTCAATG	960
Db	901	CACCAACAATCATCTGCCGAGTCACTAGCAGGATCTGTTCTTCGACAGCCCTCAATG	960
Qy	961	GGCGAAATCCTTGTAGGCTCGGTATGATGATCAAGAGTCAAGAGTGTCTCTCT	1020
Db	961	GGCGAAATCCTTGTAGGCTCGGTATGATGATCAAGAGTCAAGAGTGTCTCTCTCT	1020
Qy	1021	GACAGTGTGACCGGACCGATCATATTTCTCCCTTACAAGGTTTCAAGCATAAATAT	1080
Db	1021	GACAGTGTGACCGGACCGATCATATTTCTCCCTTACAAGGTTTCAAGCATAAATAT	1080
Qy	1081	CTACAGGAGATGCTTGATTTGTAGACATTTTGACACCTTTTACAAGTTTCTTGTCTGGGT	1140
Db	1081	CTACAGGAGATGCTTGATTTGTAGACATTTTGACACCTTTTACAAGTTTCTTGTCTGGGT	1140
Qy	1141	CACAGTCTTGCGTATCTTGTATACCTTGCGTACGAACTCAACTCAAGGTTGGAGATGT	1200
Db	1141	CACAGTCTTGCGTATCTTGTATACCTTGCGTACGAACTCAACTCAAGGTTGGAGATGT	1200
Qy	1201	GGTGGTGTGCTGGCGGTACAGTCTTGTCTTCAATTAATTAATTAATTAATTAATTAATTAAT	1260
Db	1201	GGTGGTGTGCTGGCGGTACAGTCTTGTCTTCAATTAATTAATTAATTAATTAATTAATTAAT	1260
Qy	1261	ATTCACAACCTACGATTTCTGGTGTCTTCTGATGATGTGCGTGTGCTATCGG	1320
Db	1261	ATTCACAACCTACGATTTCTGGTGTCTTCTGATGATGTGCGTGTGCTATCGG	1320



/db\_xref="gi:3252836"  
/translation="MNDGVNNGSESEPEKMEPRRSQISLPANIKATMDNDML  
EAPPEPMNSNTEPMSCTATSVGLGNWRRPFLAYONGGAFVLPYIVLLVKG  
PVYLYECVLOQFSSRSNYSKIMISPAKGGVQAQVCGGILSYVYVGLFLYAM  
SFOATLFWALICQPEWENCVPDEPLAAGVSNITNGSSAFLEFLVLOOSDIEGL  
GAPLWYLVLCFLAMLVGVNVAAGVSSKAAYFLAFLEPYVMILFTLVYLPAT  
DGLLEFVTPMVKLLEGLVMAITQVFEFLSVCTGATIMFSSNPKFNVRDALIV  
TTLDTFSLSGCTFGLIGNLAVELNSDGDVIGSGSLAISIPDKATKFDALIV  
FSVLEFLMSVGLGSSVALSTLNVDNAPRPVYVMSALSCGFLGLVYCTP  
GGOFLELVHGGTFLYLCALSELNAGVWITGLNCLDIEFMKGTGGYWRLCW  
GIYPGMMITVFYIALISLENLVFGDFTYYPVAGVAGYMLFLGLVLPIDIVLY  
KYRGNRRETVKAFHSKPSWGRPRRLREWMQFRIEAKALROKMNFSRVHLMYSI  
TGARRRINK"

BASE COUNT 770 a 607 c 610 g 894 t

ORIGIN

Query Match 84.7% Score 2423.4; DB 3; Length 2881;  
Best Local Similarity 91.6%; Pred. No. 0;  
Matches 2627; Conservative 0; Mismatches 221; Indels 21; Gaps 5;

QY 9 GTTACTTGTGAGGAACTGTTGGCGGTGTCGATGATTTCTGAACACAAATACCT 68  
DB 11 GTTACTTGTGAGGAACTGTTGGCGGTGTCGATGATTTCTGAACACAAATACCT 70  
QY 69 AACACATTCGACAGTGTGATTTGTGACAAAATGATGACGCCAAGTAAACGGCGTT 128  
DB 71 AACACATTCGACAGTGTGATTTGTGACAAAATGATGACGCCAAGTAAACGGCGCT 130  
QY 129 TCGAGTCGCGAGCCCAAGATGAAACCAAGATCATCGCAATTAAGTTACCTCCAG 188  
DB 131 TCGAGTCGCGAGCCCAAGATGAAACCAAGATCATCGCAATTAAGTTACCTCCAG 190  
QY 189 CAATTAATTAAGCGGCTCTGATTAATTAATGATGACAGGACTTAGAGCTGACCGCAG 248  
DB 191 CAATTAATTAAGCGGCTCTGATTAATTAATGATGACAGGACTTAGAGCTGACCGCAG 250  
QY 249 AACGATGATGATGTCACAAACAAATGATGATGTCGATGTCGATGTCGATGTCG 308  
DB 251 AACGATGATGATGTCACAAACAAATGATGATGTCGATGTCGATGTCGATGTCG 310  
QY 309 GTTGGGTAAAGTGGGGTCCCTTTCATCGGTCACAGAAATGAGAGAGTCTTCC 368  
DB 311 GTTGGGTAAAGTGGGGTCCCTTTCATCGGTCACAGAAATGAGAGAGTCTTCC 370  
QY 369 TGTGTCATGATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCG 428  
DB 371 TGTGTCATGATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCG 430  
QY 429 TCTCGGCAATTAAGTTCAAGAACTCTGTTAAAGTTGGTCAATTTACCGGCAATGA 488  
DB 431 TCTCGGCAATTAAGTTCAAGAACTCTGTTAAAGTTGGTCAATTTACCGGCAATGA 490  
QY 489 AAGTACTGATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCG 548  
DB 491 AAGTACTGATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCG 550  
QY 549 TCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 608  
DB 551 TCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 610  
QY 609 TTTGTGACCTGAGTGGAGAACCTGCTACCAATCCACACTTGTGTCATGTCGTA 668  
DB 611 TTTGTGACCTGAGTGGAGAACCTGCTACCAATCCACACTTGTGTCATGTCGTA 670  
QY 669 ACAACATGCAATGTCACGACGATGTCACACTTCTTCTTGTGAGAACGTTGTCAC 728  
DB 671 ACAACATGCAATGTCACGACGATGTCACACTTCTTCTTGTGAGAACGTTGTCAC 730  
QY 729 AAGCGATGAATTTGAAGGAGTCTGGGGCCCATCTGTCGATGTCGTCGTCGTCAT 788  
DB 731 AAGCGATGAATTTGAAGGAGTCTGGGGCCCATCTGTCGATGTCGTCGTCGTCAT 790  
QY 789 TCAATGCAATGCTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 848

DB 791 TCAATGCAATGCTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 850  
QY 849 CGGCTACTTCTCGCGCTCTCCCATACGTTGTCATATACCTTTATTCATCCACAA 908  
DB 851 CGGCTACTTCTCGCGCTCTCCCATACGTTGTCATATACCTTTATTCATCCACAA 910  
QY 909 TCATCTCGCGCGCTCTCCCATACGTTGTCATATACCTTTATTCATCCACAA 968  
DB 911 TCATCTCGCGCGCTCTCCCATACGTTGTCATATACCTTTATTCATCCACAA 970  
QY 969 TCTTTGACCTGCGGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1028  
DB 971 TCTTTGACCTGCGGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1030  
QY 1029 GCACCGGACCATATATATATGTCCTCTCTTCAACAGGTTTCAGACATATATATAC 1088  
DB 1031 GCACCGGACCATATATATATGTCCTCTCTTCAACAGGTTTCAGACATATATATAC 1090  
QY 1089 ATGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1148  
DB 1091 ATGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1150  
QY 1149 TCGGATATCTTGGTATACCTCGGTAGCAACTCAGAGTGGAGATGTCGTCGTCG 1208  
DB 1151 TCGGATATCTTGGTATACCTCGGTAGCAACTCAGAGTGGAGATGTCGTCGTCG 1210  
QY 1209 CTGCGGTAGCAACTCCTGTCATTTATTCATGATGATGATGATGATGATGATGAT 1268  
DB 1211 CTGCGGTAGCAACTCCTGTCATTTATTCATGATGATGATGATGATGATGATGAT 1270  
QY 1269 CTGCGGTAGCAACTCCTGTCATTTATTCATGATGATGATGATGATGATGATGAT 1328  
DB 1271 CTGCGGTAGCAACTCCTGTCATTTATTCATGATGATGATGATGATGATGATGAT 1330  
QY 1329 TGGCTCTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1388  
DB 1331 TGGCTCTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1390  
QY 1389 TCTACATGTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1448  
DB 1391 TCTACATGTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1450  
QY 1449 CGGCTGACAAATATATCTTGAAGTGTGATGATGATGATGATGATGATGATGAT 1508  
DB 1451 CGGCTGACAAATATATCTTGAAGTGTGATGATGATGATGATGATGATGATGAT 1510  
QY 1509 TCTGCGGCAATTTCTGACCTCGCAGCGTGTCTGATTTATGATGATGATGATGAT 1568  
DB 1511 TCTGCGGCAATTTCTGACCTCGCAGCGTGTCTGATTTATGATGATGATGATGAT 1570  
QY 1569 TAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1628  
DB 1571 TAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1630  
QY 1629 TAATACGTCGTCATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1688  
DB 1631 TAATACGTCGTCATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1690  
QY 1689 TGTGTCGAGGACAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1748  
DB 1691 TGTGTCGAGGACAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1750  
QY 1749 TATTTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1808  
DB 1751 TATTTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1810  
QY 1809 CCGGACATCTGAGGACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1868  
DB 1811 CCGGACATCTGAGGACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1870  
QY 1869 GCTCGCGAGAGACGTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1928



Db	2322	TGGAAACAGCAATCGAGTTCTCTCATGTCTCATGCACTGACGCTGGGTCAT	2381
QY	320	GTGTGGCGGTTCCTTTTCATCGGTAACGAATGAGAGAGTCTTCTGGTGCATAC	379
Db	2382	GTGTGAGAGTTCCCTTACCGCACTAGAGAAACGGAGAGAGCCTTGTGTAATCC	2441
QY	380	GTATGCTCTTTTACTGTGCGCAAGCCTGTGTACTTATGATGTGTCTGGACAA	439
Db	2442	CTGATGTCTCTGATTTCTAGTGGCAAGCCCGGTACTACTTGGAAATCTGTCCG	2501
QY	440	TTTCAGTTCAAGAACTCTGTAAAGTTGGTCAATTTACCGGGCATGAAAGTA	499
Db	2502	TTTCTGAGTCGGGGCAGTGTGAAGGTCTACAGCTTTACCCATATATAGAGCA	2561
QY	500	TACGCTAAGCTGCGGCTGCGGTTACATCTGTCTTACTACGTTGATGTGTCTC	559
Db	2562	TATGGAACAAGTCTGTGCCCAAGGACATGTTTACACCTACTACGCAAGTAA	2621
QY	560	TGCTGTATATATATAGCTATGAGCTTCCAGGGCACTTTCATGGGCTATTTG	619
Db	2622	ACCCATACGCTACTTGTGAGACTCTTTTACCCGACCTTGGGTGAGCTACTGC	2681
QY	620	GAGTGGG-----AGACTGCTACCTTCAGATCCAAACTGTCTCATAGTC	667
Db	2682	GAGTGGGTCAGAAATGCGCTGACATCGGAGACGAGAGGCGACCAAGTTTG	2741
QY	668	AACAACATCAACAATGTAACCAAGCAAGTCTCAACTTCTTTGAGAACGTTCT	727
Db	2742	GCAGGCTCGGAGTCCGAACCACTCGGCGGAGTTCTATTTACAGAACATATCT	2801
QY	728	CAAAAGCATGTAATTTGAAGAGAGTTCGCTGCCCACTGTGTTGGTTGGTCTA	787
Db	2802	GAGAAAGGAGCATGATGATGAGGATCGGATATCCAGCTGAGCCTGGCCTGG	2861
QY	788	TTTATTCGATGCTCATGTGTGTGGAGTCTCGCCGAGAGTCAAGAGTTCCGG	847
Db	2862	GCTGTGGCTGTGATGTATCGCGGGGATTTATTTCAAGGSTGTGAAGATGCG	2921
QY	848	GCGGCTACTTCTCGCGCTCTTCCCATAGTGTTCATGATCACTTTATTCATCA	907
Db	2922	GCTCTCTACTTCTTCTGCTCTTCTTCCCTACGTGGTGTGCTTGTGTGTCGG	2981
QY	908	ATCATCTGCCCGGTGCTACTAGAGGACATCTGTCTTCTGTCACAGCTTCA	967
Db	2982	CTACACTTTCGGGGGCTTCGACGAGGGGTGTACTTCTCGGGCCGCAATGCA	3041
QY	968	CTCCTTAGCTCGGCTATGTGTACTACAGACTACGCAAGTCTTCTCTGACAG	1027
Db	3042	CTTCTGGAAACGCAAGTGTGTAGCGCGCTTCCACCAAGGTCTTCTACAGC	3101
QY	1028	TGCACCGACCAATCATCATCTTCTCTTACCAAGSTTTACAGACATATATCT	1087
Db	3102	TGCTTGGGAATCATCATATGTACGCTGTGACCAACGCTTCGGCCCAACATTT	3161
QY	1088	GATGCTTGATTTGTACGACTTTGGACACTTTTCAAGTTCTTGTCTGGGTCA	1147
Db	3162	GATGCAATATCTGACACACGCTGACACCTTCACTCCCTACGTGCGGTGTAT	3221
QY	1148	TTGCGTATCTTGTAACTCGGCTAGAGACTCAACTCAAGTGGGATGTGTGCG	1207
Db	3222	TTTGGGATTTCTGGCAACCTGGCTGTAGAAACAACCAACGCAATCGTACTG	3281
QY	1208	GCTGCGGATACCACTTCTGTTTCAATTTACCCATATCCATTTGGCA--A	1264
Db	3282	AATGAGAGTCCGGGCTGTGCTTATATCTTACCGGACGCCATTGCCAAGTT	3341
QY	1265	CAACCTAAGCTATTTCTGGGTGTCTTCTCTATGATGTGCGTGGGTATGCG	1324
Db	3342	CTGCCGAGCATTTCTTCCGTGCTTCTTCTCATGCTCTTCTCTGGGCAATG	3401
QY	1325	TCCGTGCTCTGCTATGCACTTTCACACATTTGGCATATGAGCGCTTCCAG	1384

Db	3402	AATGTGGGCGATGGCCCTCTGCATGTCACCGCTGATCAAGATACAGTGGAGACCGTGAAG	3461		
QY	1385	ACCGGTACATATGACAGCCATGACACCTGTCTTTCGCGTTCCCTCTGGACTTGTACTGC	1444		
Db	3462	AACGTGACATGTGGTGGTGGTATGACCATTTGAGCTACTTCTGGGCGCTCTGTACATC	3522		
QY	1445	ACACCGGGTGAGCAATATATTTCTTGAGCTTTGATATCACTACGGTGAACATTCCTTGG	1504		
Db	3522	ACACCGGGTGAGCGAGTTCCTGCTCAACCTGTCGATCACTTTGGCGTCACTGCTGGCC	3581		
QY	1505	CTTTTCTGGCCATTTTCTGTAACCTCCAGGCGTGTCTTGATTTATGATTTGAGAACTGC	1564		
Db	3582	CTGGTCTGGCCATCTTTCAGGTGTGACCATTTGCTGGATCTACGGTGTGAAGGACCTC	3641		
QY	1565	TGCTAGACATTTGAGTTCATGTTGGGTAAAAAGACTGGTGGCTTACCTGGCGTCTGGTG	1624		
Db	3642	TGCGGTGACGTGAGTTCATGATTTGGCATCAAGACTTGGCTGACTATTCGATCTGCTGG	3701		
QY	1625	GCGCTAATCATCTCTGCTATTAATGACGACTGTGTTCTTACGCTCTTCTGCGCTTAAC	1684		
Db	3702	GCGGTCTGTCACCTCTGCTATGCTACACATTCATCTCATCTACACTTGGTCTGTACAG	3761		
QY	1685	AACCTGTGTTGGAGACACACTAGCTAATACCGACTGCT	1723		
Db	3762	CCCCTCAACTACAGAGATTACACTCAACCAATGCGGTGT	3800		
RESULT 4	AXI91697	2473 bp	DNA	linear	PAT 15-AUG-2001
LOCUS	Sequence 3 from Patent WO0149848.				
DEFINITION	AXI91697				
ACCESSION	AXI91697.1	GI:15209877			
VERSION					
KEYWORDS	fruit fly.				
SOURCE	Drosophila melanogaster				
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 2473)				
AUTHORS	Kelleman, K.A., Keegan, K.P., Ebers, A.J. and Norpey, J.				
TITLE	Nucleic acids and polypeptides of drosophila melanogaster snf sod1 sodium neurotransmitter symporter family cell surface receptors and methods of use				
JOURNAL	Patent: WO 0149848-A 3 12-JUL-2001;				
FEATURES	Genoptera, LLC (US)				
source	Location/Qualifiers				
	1..2473				
	/organism="Drosophila melanogaster"				
	/db_xref="taxon:7227"				
BASE COUNT	625 a 593 c 627 g 628 t				
ORIGIN					
Query Match	13.7%;	Score 391.8;	DB 6;	Length 2473;	
Best Local Similarity	53.2%;	Pred. No. 8.6e-75;			
Matches	951; Conservative	0; Mismatches 767;	Indels 69;	Gaps 3;	
QY	218	GATACACGGACTTAGAGGCTGAACGCCACGAACGTAATGATGCTCAACAACATTGAA	277		
Db	327	GAATAAAACAGATGCGGAGAGCCGACGCGGAGCGACCAATTGGGCACTGTGAGTCCGTGAG	386		
QY	278	TTTCTGATGTCCTCATCGCTACACATCCGTGGTTTGGGTAACTGTGGCGCTTCCCTTC	337		
Db	387	TTTCTAATGTCTCGCATATGCGTGTCCGTGGATTTGGCAACGTCGTGAGTTCCTCCGTTCC	446		
QY	338	ATCCGCTACAGATGAGGAGGTGCTTCTCTGTCGATACCTATCGTTCTTTTACTT	397		
Db	447	ACGGCTATGAGATGAGGAGGCGCTTCTTATATACCTCATCATATGATGCTCTTTCTG	506		
QY	398	GTGGGCAAGCTGTGTACTTATAGAGTGTCTCTGGACATTCAGTTCAAGAACTCT	457		
Db	507	ATCGGCAACCCATGTAATTTTGGAGATGATCATGAGCCAGTTCAAGACCAGGGAACG	566		



QY 458 GTTAAAGTTTGGTCAATTTTCCGGCCATGAAAGTACTGATACGCTACGCTGCCGC 517  
 DB 567 GTGAAGATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGC 626  
 QY 518 TGGGGTACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 577  
 DB 627 ACCATCTGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 686  
 QY 578 ATGAGCTTCCAGGCGCTCTTCCATGAGGCTATTTGTACAGCTGAGTGGAGAACTGGCTA 637  
 DB 687 GTGTCTCTCAATCGAATGCGGTGGTCTTACTGTCGCGATGAGTGGACCAATTTGGGTG 746  
 QY 638 CCCGC----- 642  
 DB 747 AACTCGAGCCCGAGAGTATGTGGACAATCTGCTTACGGGCGTATCTTGGCCAAATGAA 806  
 QY 643 -----AGATCCAAACCTTGTGATCATCAACCAATCAACCAATGATGACAGAGT 694  
 DB 807 TCGGCCAAGAAATCTCACGGGCAATGTGGCCAGCATGAGACAGAAATCCACAGACAGC 866  
 QY 695 GCTCACTACTTCTTTTGAAGACAGTCTCCACAAAGCATGGAATGGAAGAGTCTC 754  
 DB 867 TCCGAACCTATTTCTTGAATGTGATGACAGAGAGAGCTGAGCATCTGGAGCGGCTT 926  
 QY 755 GGTGCCCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 814  
 DB 927 GGTGATCCCGATTGGAAGCTGACACTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 986  
 QY 815 GTGCTGCGCGAGAGAGTTCGGGCAAGCGGCTTACTTCTGCGCTCTTCCCA 874  
 DB 987 GTGATCATGCGAGGGGTGAAGAGTTCGGGCAAGCGGCTTACTTCTGCGCTCTTCCCA 1046  
 QY 875 TACCTGTCAATGATCACTTTATTCATACCAACCAATCAATCTGCGCGGCTGCTGCTGCTG 934  
 DB 1047 TACCTGTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106  
 QY 935 ATCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 994  
 DB 1107 ATCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1166  
 QY 995 GCACTCAGCAGAGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1054  
 DB 1167 GCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1226  
 QY 1055 TCTTACAACGGTTTACAGCAATATATATACAGGAGTGTGATGCTGCTGCTGCTGCTGCTG 1114  
 DB 1227 TCTTACAACGGTTTACAGCAATATATATACAGGAGTGTGATGCTGCTGCTGCTGCTGCTG 1286  
 QY 1115 ACCCTTACAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1174  
 DB 1287 ACGCTGACACCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1346  
 QY 1175 GAACCTCACTCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1234  
 DB 1347 AATCTGCAATCGAAACATCAGAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1406  
 QY 1235 TCAATACCGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1291  
 DB 1407 TCGTATCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1466  
 QY 1292 TTTCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1351  
 DB 1467 TTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1526  
 QY 1332 ACATTTGGCGATGAGCGCTTCCACAGCTTACCCACCGCTTACATGCTGACGATGCTGCT 1411  
 DB 1527 ACCATCATTTGGCGATGAGCGCTTCCACAGCTTACCCACCGCTTACATGCTGACGATGCT 1586  
 QY 1412 TCTTCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1471  
 DB 1587 GTGTGCGGTTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1646  
 QY 1472 CTGTGATCACTACGCTGGAACATCTCTGTGCTTCTGCGCCATTTCTGCACTGCCA 1531

DB 1647 TTGGTGGACTTCTATGCGGACCATATGCTGCTTACTTCTGGCCATTTTGCAGAACTGCT 1706  
 QY 1532 GCGGTGTTCTGATTTATGATGATGAGAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1591  
 DB 1707 GGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1766  
 QY 1592 AAAAAGCTGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1651  
 DB 1767 CGCGAGTTTCACTGATCTGCGGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1826  
 QY 1652 ACTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1711  
 DB 1827 ATCATATTCAATATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1883  
 QY 1712 TACCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1771  
 DB 1884 TTTCCCGAAGCGGCAACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1943  
 QY 1772 CCATTTGGAATGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1831  
 DB 1944 CCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2003  
 QY 1832 AAGAAAGCTTCCACTCCAAACCTCATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1891  
 DB 2004 AAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2063  
 QY 1892 TGATGCACTTCAAGGCTGAGAGCAAGCTTGTGACAAAAGATGAA 1938  
 DB 2064 TGGGTGATCTTCAAGATTCAGAAAGCAAGCCCAAGAGCCACGACGAA 2110

## RESULT 5

AY121662

LOCUS

AY121662

DEFINITION

Drosophila melanogaster full insert cDNA.

ACCESSION

AY121662.1

VERSION

GI:21464371

KEYWORDS

FLI.CDNA.

SOURCE

fruit fly

ORGANISM

Drosophila melanogaster

REFERENCE

1 (bases 1 to 2497)

AUTHORS

Stapleton, M., Brokstein, P., Hong, L., Aghayani, A., Carlson, J.,

TITLE

Direct Submission

JOURNAL

Submitted (13-JUN-2002) Berkeley Drosophila Genome Project,

COMMENT

Lawrence Berkeley National Laboratory, One Cyclotron Road,

Berkeley, CA 94720, USA

Sequence submitted by:

Berkeley Drosophila Genome Project

Lawrence Berkeley National Laboratory

Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to

sequence clones from Drosophila Gene Collection 1 (Rubin et al.,

Science 2000). The sequence has been subjected to integrity checks

for sequence accuracy, presence of a polyA tail and contiguity  
 within 100 kb in the genome. Thus we believe the sequence to  
 reflect accurately this particular cDNA clone. However, there are  
 artifacts associated with the generation of cDNA clones that may  
 have not been detected in our initial analyses such as internal  
 priming, priming from contaminating genomic DNA, retained introns  
 due to reverse transcription of unsplined precursor RNAs, and  
 reverse transcriptase errors that result in single base changes.  
 For further information about this sequence, including its location  
 and relationship to other sequences, please visit our Web site  
 (<http://fruitfly.berkeley.edu>) or send email to

```
cdnaefruitfly.berkeley.edu.
Location/Qualifiers
1..2497
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db.xref="taxon:7227"
1..2497
/gene="CG3252"
/db.xref="FLYBASE:FBgn0029762"
280..2205
/gene="CG3252"
/note="Longest ORF"
/codon_start=1
/product="RE10560p"
/protein_id="AAM51989.1"
/db.xref="GI:21464372"
/db.xref="FLYBASE:FBgn0029762"
/translation="MELKGVQPSNGSSNGSGNGATNAASTEKTDIAEKPTAERTMNGNG
LEFLMSTISVSVGLGNWRPEFTAYENGGAFLPIYILVFLGKPMYLEMNGOPT
SOGKTIWSVYVPGVYGQAFETICIIISYSLALTLVLYLFSQSLPMSYCD
EMNCVNSRQEVYDNLITGSLANESARMLSGIVANDEKLOSSELYELNVYKE
KLIDSDGVDPMDKLTALFVAVVYFLVIMRGYKSSGKRAYFLALPPYVLELIR
AVTLEGARDGLFELBPQWELLNPTVKRAVVOCPSLAVSGSPITMPSYRFBHG
IYRDAMVTLTDLTSLSGITTFPAILGNLAHNIQENIDVRSGLAFISPDAL
SKRDVQQLSVLEFEMLVYLGISVALDSTLVTTICDQKGMKRWKVALTTSVGF
LMGLVYTPGQWLTLLVDFYGGYVFLAIFELAGIYVWYGLONCDIEMCNR
VSLVRVMSRFPVMMILFIYSMTVIEIKYSELFEPAANAGLWLFALICAOP
LMGLMYISRHDQYMKSLKSLKPSDRMGPPANDEIRRENVIFKNQAAQRAIQKGS
KLGFPMKRVNFCGSNK"
BASE COUNT      625 a      604 c      636 g      632 t
ORIGIN
Query Match      13.7%; Score 391.8; DB 3; Length 2497;
Best Local Similarity 53.2%; Pred. No. 8.6e-75;
Matches 951; Conservative 0; Mismatches 767; Indels 69; gaps 3;
QY 218 GATGACGCGACTTAGAGCGCTGAACCGCCAGAACGATGATGATGTCACAAACATGGA 277
DB 358 GAAAAACACATGCGCGAAGCCGCGGAGCGGACCCCAATTGGGCAATGGCCTGGAG 417
QY 278 TTCTTATGTCCTGACATCGCTACATCCGTCGTTTGGATGAGCGGTGGCCCTTCC 337
DB 418 TTCTTATGTCCTGACATCGGTGTCGTGGGANTGGGCAACGCTGAGAGTTCCCGTTC 477
QY 338 ATGCGGACAGAAATGAGAGAGTCTTCTGTCGTCATGCTCATGCTTCTTAACT 397
DB 478 ACGGCTATGAGATGAGAGGCGCTTCTTATACCTTACATCATAGTCTCTTCTG 537
QY 398 GTGCGCAAGCCTGTGACTTACTTAGAGTGTCTCTGGAGCATTCAGTTCAANAAC 457
DB 538 ATGGGCAACCCATGTCATTTTGGAGATGATCATGGGCCAGTTCAAGGCCAGGAA 597
QY 458 GTTAAAGTTGGTCAATTTACCGGGCCATGAAGATGATGATACCTCAAGTCGCGG 517
DB 598 GTCAAAATTTGGTGGTGTGGTGGGATTTGTGGCGGTGGGCTATGGCCAGGCTTC 657
QY 518 TGGGTTAAATCTGTCCTTACTACGTCGTGATGTCGTGTCGTCTGTTACTTA 577
DB 658 ACCATCTGCATATCTGTCATATTCCTGCTGTGGCGTGAACCTTCAATCTCT 717
QY 578 ATGAGCTTCAGGCGCACTCTTCATGAGGCTATTTGACAGCTTGGGAGGAACTG 637
DB 718 GTGTCTTCCAAATCGGAATTTGCCGTGTCCTACTGTGCGGATGATGAGCCAA 777
QY 638 CCCTC----- 642
DB 778 AACTCGAGCCGAGAGATATGGAACAATCTGCTTACGGGCGTATCTTGGCAATGA 837
QY 643 -----AGATCCAACTTGTGTCATAGTCACAAATCATCAACCAATGATACCA 897
DB 838 TGGGCGAAGAAATCTCAGCGCATTTGGCCAAAGATGAGACAGAGAAACTCCAG 897
QY 695 GGTCAACTCTACTTTTGGAAACAGTTCTCAACAACAAGGATGATGATGAAGAGG 754
```

```
DB 898 TCCGAACCTCTATTTCCTGAAATGTAATGATACAGAGAACTGGACATCTTGGACGG 957
QY 755 GGTGCCCCCATCTGGTACTTGGTGTGTCTATTATCCAGATGAGCATGTGTTCGA 814
DB 958 GGTATCCCGATGGAAGCTGACACTTGGCCCTGTGTCGCTGGGTGGTGCATCTTCTG 1017
QY 815 GTCTGCGCCGAGAGATCAAGAGTTCGGCAAGCGGCTTACTTCTCGCGCTTCCCA 874
DB 1018 GTGATCATGGAGGAGTGAAGAGTCCGGAAGCGGCGCTTACTTCTCGCGCTTCC 1077
QY 875 TACGTTCTATGATCACTTATTATTCATACCAATCATCTGCGCGGCTTACTGACGG 934
DB 1078 TACGTGTGCTCTTCTGCTGCTGCTGATCCGCTGCTTACTTGGAGGAGACCGCAT 1137
QY 935 ATCTGTTCTTCTGTCACGCTCAATGCGCAAACTCTTGAAGCTGGTGTATGATCTCA 994
DB 1138 ATCTGTTCTTCTGAGCCACAGATGGGAGTGAATCTTGAACCCACCGCTTGGAA 1197
QY 995 GCAGTCACGAAGTCTCTCTCTGACAGTGTGACCGGACCGATCATATGTTCTC 1054
DB 1198 GCGGTGTGCAATGCTTCTCTGCTGCGCGGCTGCGGACCATATATCATGTTGCGC 1257
QY 1055 TCTTACAAAGCTTTCAGACATTAATATCTACAGGATGCTGATGTTGATGACATTGGAC 1114
DB 1258 TCGTACAAATGATTCGATTCAGCAATCTATAGGATGCCATGATCGTAAACACATG 1317
QY 1115 ACCTTTACAAATGTTCTTCTGCTGCTGACAGATCTTGGTATCTTGGTAACTCG 1174
DB 1318 ACGCTGACCAAGCTCTCTGGGAGTATAGCATATTTGCGATCTTGGGCAATCTGG 1377
QY 1175 GAATCAACATCAGAGTGGAGATGTGTGCGGTGCGGATCCAGTCTGCTTCAAT 1234
DB 1378 AATCTGCAATCGAATAAATCATCAGATGTGTGTCGAATGTCACGCGGATTTGGCTT 1437
QY 1235 TCAATACCTGATCCATTTGCCAA--AATTTCAACCTTACGATTTCTGCTGCTGCT 1291
DB 1438 TCGTATCCGATCCATATCAAAATTCACAGGCGGTGCGGAGCTTTCTCGGTTCTG 1497
QY 1292 TTCTGATATATGTCGTCGCTGCTGATGCTGCTATCCGTGCTGCTATGACATTCA 1351
DB 1498 TTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1557
QY 1352 ACATTTGCGATGACGCTTCCACGCTGATACCCAGCTGATATGTCACAGATGAC 1411
DB 1558 ACCATATTTGCGATGATGATCAAGGATGGAAGTACTGGAAGTGGCCTGACAC 1617
QY 1412 TCTTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1471
DB 1618 GTGTGCGGTTTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1677
QY 1472 CTGTGATGATCACTACAGTGAACATCTCTTGTGCTTGTGCGGCTTCTGAC 1531
DB 1678 TTGTGTGACTTCTATGCGGAACTATGTGTGCTTCTTGTGCTTCTTGTGCTT 1737
QY 1532 GCGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1591
DB 1738 GGCATGCTGTGGTATGCTGCAAACTTTTGGATGATGATGATGATGATGATG 1797
QY 1592 AAAAAAGCTGCTTACTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1651
DB 1798 GCGCGAGTTTCACTGATGCTGCGAGTATGCTGCTGCTGCTGCTGCTGCTGCTG 1857
QY 1652 ACTGTTCTTCAAGCTTCTTCCGCTTCAACACCTGCTGCTGCTGCTGCTGCTG 1711
DB 1858 ATCATATTCATATCTGATGATGATGATGATGATGATGATGATGATGATGATG 1914
QY 1712 TACCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1771
DB 1915 TTCCCGGAAAGCGGCAACATTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1974
QY 1772 CCAATTTGGAATGATTTTCTTGTGACAAATACCTGCTGCTGCTGCTGCTGCTG 1831
```

Db 1975 CCATGTGGGGTCTTTGGTACATCTCGGCCATCCACAAGGACCTACTGGAAGTCTCTG 2034  
 QY 1832 AAGAAACCTTCCATCCAAACCTCTATGGGGTCCCGCTCGCCGAGAGAGCTAGAGAA 1891  
 Db 2035 AAGGCTCGCTGGAAGCCAGGATCGCTGGGCTCTCCCAATCCGAAATTCGGCGGAA 2094  
 QY 1892 TGGATGAGTTCAGGCTGAAGGAAAGCTCTTACAAAAGATGAA 1938  
 Db 2095 TGGGTGATCTTCAGAAATCAGAAAGGCAAGCCCAAGAGCCGACAGAA 2141  
 RESULT 6  
 AX191701 2030 bp DNA linear PAT 15-AUG-2001  
 LOCUS Sequence 7 from Patent WO0149848.  
 DEFINITION AX191701  
 ACCESSION AX191701.1 GI:15209879  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 fruit fly.  
 Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidae; Drosophilidae; Drosophila.  
 1 (bases 1 to 2030)  
 Kelleman, K.A., Keegan, K.P., Ebens, A.J. and Torpey, J.  
 Nucleic acids and polypeptides of drosophila melanogaster snf sodi  
 sodium neurotransmitter symporter family cell surface receptors and  
 methods of use  
 Patent: WO 0149848-A 7 12-JUL-2001;  
 JOURLAL Genoptera, LLC (US)  
 FEATURES  
 source  
 1. 2030  
 Location/Qualifiers  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 BASE COUNT 455 a 470 c 507 g 598 t  
 ORIGIN  
 Query Match 13.4%; Score 383.4; DB 6; Length 2030;  
 Best Local Similarity 55.3%; Pred. No. 5.8e-73;  
 Matches 788; Conservative 0; Mismatches 631; Indels 6; Gaps 2;  
 QY 260 TGGTCCAAACATTAATTTCTTGATGTCCTGATCGCTACATCCGCGGTTGGGTAC 319  
 Db 108 TGGGACATGCTGCTGATGTCCTGATGTCCTGATGTCCTGATGTCCTGATGTCCTG 167  
 QY 320 GTGTGGCGGTTCCCTTTATCGCGGATGAGATGAGAGGTCTTCTGTCGCATAC 379  
 Db 168 GTGTGGCGGTTCCCTTTATCGCGGATGAGAGGTCTTCTGTCGCATAC 227  
 QY 380 GTATCGTCTTCTTCTTCTGTCGCAAGCTGTGATCTTACTAGTGTGTCCTGACAA 439  
 Db 228 CTCGTGCTCTTCTTCTGTCGCAAGCTGTGATCTTACTAGTGTGTCCTGACAA 287  
 QY 440 TTCACTTCAAGAACTGTTAAAGTTGGTCAATTTCAACGCGCATGAAAGTACTGA 499  
 Db 288 TTCTCAGATCGTGGATCGTGGATCGTGGATCGTGGATCGTGGATCGTGGATCG 347  
 QY 500 TACGCTCAAGTCCGGGCTGCTGATCATCTGCTTACTAGTGTGATCTGCTGCTC 559  
 Db 348 TATGCTCAGCTATTTGGCCCTAGGCGCTTTGGCCAGCTATTTAGCTTCACTGATG 407  
 QY 560 TGTCTATTTACTTACTAGCTATGAGCTTCCAGGCACTTCCATGAGGCTATTTGTCAG 619  
 Db 408 ACTCTGCGATCTTTTTCATCTCTGATCGATGAGGATTCCTCGGAGCTTCTGTGCG 467  
 QY 620 GAGTGGAGAAC--TGCTACCTTCAGATCCAACTGTCGATCACTCAACAATC 676  
 Db 468 GAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 527  
 QY 677 ACCAATGATACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 736  
 Db 528 CATACTTCAAGCTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 587

QY 737 GAAATGAGAGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 796  
 Db 588 TCCCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 647  
 QY 797 TGGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 856  
 Db 648 TGGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 707  
 QY 857 TTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 916  
 Db 708 GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767  
 QY 917 CCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976  
 Db 768 CCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 827  
 QY 977 CTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036  
 Db 828 CCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 887  
 QY 1037 CCGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1096  
 Db 888 GTGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 947  
 QY 1097 ATGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1156  
 Db 948 ATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1007  
 QY 1157 CTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1216  
 Db 1008 CTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1067  
 QY 1217 ACCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1273  
 Db 1068 CCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1127  
 QY 1274 CTATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1333  
 Db 1128 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1187  
 QY 1334 CTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1393  
 Db 1188 ATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1247  
 QY 1394 ATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1453  
 Db 1248 ATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1307  
 QY 1454 GACATATATATCTGAGCTTGTGATGATGATGATGATGATGATGATGATGATGATG 1513  
 Db 1308 GGTCAATATATATCTGAGCTTGTGATGATGATGATGATGATGATGATGATGATG 1367  
 QY 1514 GGCATTTTCAAGCTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1573  
 Db 1368 GGCATTTTCAAGCTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1427  
 QY 1574 ATTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1633  
 Db 1428 GGTGATATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1487  
 QY 1634 ACTCTGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1678  
 Db 1488 ACTCTGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1532

RESULT 7  
 AY071318 2099 bp mRNA linear INV 20-DEC-2001  
 LOCUS AY071318  
 DEFINITION Drosophila melanogaster RE34084 full length cDNA.  
 ACCESSION AY071318  
 VERSION AY071318.1 GI:17945784  
 KEYWORDS FLI.CDNA.  
 SOURCE Drosophila melanogaster.





Db 563 TATCCACTGATTAACTAGCCCCGATGCGAGTAATGTGTCACTGGAAGCAATGA 622  
 QY 648 -CAACACTTGTGATCAGTCAACAATCAACCA-----TGTACACAGACTGCTCA 699  
 Db 623 TCAGGCCCCGCAAAATTTACACAATGAAATACAAAGAGTGGGTCAACACAGTTGCGA 682  
 QY 700 ACTTACTTTTTGGAACAGTTCTCAACAAGCAGTGAATGAAAGAGGTCTGGTGC 759  
 Db 683 ATGTACTTTTGAAGAAAGTGTGACAGCAAAACCAATATGAGGAAGAAATCGGCT 742  
 QY 760 CCCCATCTGTACTGTGTGTGTCTATATATGATGATGCTGCTGTTGGAGTGT 819  
 Db 743 ACCAAATGGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 802  
 QY 820 CGCCGAGAGAGTCAAGAGTCCGCAAGCGGCTTACTTCTGCGGCTCCCATAGT 879  
 Db 803 TGTGCGGGAGTGAAGAGTGGGAAAGCATCTTACTTCTGCGGCTTCTCCGCTAGT 862  
 QY 880 TGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 939  
 Db 863 CATCATGGGT 922  
 QY 940 GTTCTGTGACGCTCAAGTGGGGAAGCTCTGAGCTGGGTATAGTATGATGATGAT 999  
 Db 923 CTACTTTATTAAGCCGAATGGGAAAGATTTTGGACCCGAAGTTGTGATGATGATGAT 982  
 QY 1000 CACGCAAGTGTCTCTCTGACAGTGTGACCGGACCGCATCATGTTCTCTCTCTTA 1059  
 Db 983 AACTCATGTTTGTACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1042  
 QY 1060 CAACGGTTTCAACATATATCTACAGAGATGCTGGATTTGTAAGCTTTGGACACTT 1119  
 Db 1043 CAACAACTGGGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1102  
 QY 1120 TACAAGTTTCTGT 1179  
 Db 1103 GACCTCTGT 1162  
 QY 1180 CAACCTGAGAGTGGGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1239  
 Db 1163 CGGAACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1222  
 QY 1240 CCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1296  
 Db 1223 TCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1282  
 QY 1297 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1356  
 Db 1283 CATGCTGT 1342  
 QY 1357 GGGGATGAGAGCTTCCACAGCTGATCCACCGCTCATGATGATGATGATGATGATGAT 1416  
 Db 1343 CATTCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1402  
 QY 1417 CGGTTTCTGT 1476  
 Db 1403 CAATTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1462  
 QY 1477 AGATCACTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1536  
 Db 1463 GGAATTTTGGGCGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1522  
 QY 1537 GTTCTGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1596  
 Db 1523 TGGATGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1582  
 QY 1597 GACTGTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1656  
 Db 1583 AGTGGGCTGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1642  
 QY 1657 GTTCTTACAGCTCTTCTGCTTAAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1716  
 Db 1643 TCGATCTA---TTCTATGCGACCTAACACCTCTAACATACAAATATGCTATATCC 1699

QY 1717 GACTGCTGTTATGTTCTTGTGATATCTGATGTTATTTTGGCATGACGTTTGCCAAT 1776  
 Db 1700 AAATTTGCTTACTCAATTTGTGTGGCTTAAGCGCCCTTGATATCTTCCAGCTGCCAT 1759  
 QY 1777 TGAATTTGGAATTTTCTTGTATCAAAATATGATGATGATGATGATGATGATGATGATGAT 1836  
 Db 1760 TTGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1819  
 QY 1837 AGCTTCCACTCCAAACCTCATGAGGTTGCCG 1869  
 Db 1820 AGCCTTACCGCCGAAAGAAATGGGAGCCAG 1852  
 RESULT 9  
 AX191703  
 LOCUS  
 DEFINITION  
 AX191703  
 ACCESSION  
 VERSION  
 AX191703.1 GI:15209880  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Fruit fly.  
 Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 2108)  
 Kellerman, K.A., Keegan, K.P., Ebens, A.J. and Torpey, J.  
 Nucleic acids and polypeptides of drosophila melanogaster snf sodi  
 sodium-neurotransmitter symporter family cell surface receptors and  
 methods of use  
 Patent: WO 0149848-A 9 12-JUL-2001;  
 Genoptera, LLC (US)  
 FEATURES  
 source  
 location/Qualifiers  
 1..2108  
 /organism="Drosophila melanogaster"  
 /db-xref="taxon:7227"  
 BASE COUNT 524 a 501 c 518 g 565 t  
 ORIGIN  
 Query Match 13.2% Score 378.6; DB 6; Length 2108;  
 Best Local Similarity 53.2% Pred. No. 6, 4e-72;  
 Matches 976; Conservative 0; Mismatches 799; Indels 58; Gaps 6;  
 QY 88 TATTGTGCAAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 147  
 Db 38 TACAGTGGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 97  
 QY 148 GATGGAACCAAAAGATCATGCGAATTAAGTTTACCTCGAGCAAAATATTAAGCGGCTCT 207  
 Db 98 GACGAGACGAGTATGCGGCGACAGATGAGATTTTGACG---AACAGTCCAGGCCCTGGG 153  
 QY 208 GATATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 267  
 Db 154 AGATTAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 213  
 QY 268 CAACATTAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327  
 Db 214 GGGCGTGGAGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 273  
 QY 328 GTTCCCTTATGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387  
 Db 274 ATTCCCTTATGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 333  
 QY 388 TCTTACTTGTGCGAAGCCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447  
 Db 334 CCGTTCGATTTGGGAAACCGATTTACTATGGAATGATTAATGGCAATTTTCCAG 393  
 QY 448 AAGAATCTGTTAAAGTTTGTGCAATTTCAACGGGCAAGGATGATGATGATGATGATGATGATGAT 507  
 Db 394 TCGGGGCTGGTTAAAGTTTGTGATTTGTGCCCGGATGAAAGGTGTTGGAGCTGGCCA 453  
 QY 508 ACCTGCGGCTGGGTTTACATCTGTCTTACATGATGATGATGATGATGATGATGATGATGATGAT 567

Db 454 GCGTTTCAGAGTGTTCATGCTCAGACACTTACTATGCCCGCTGGTAGCCATTAATTTGGCCG 513  
 QY 568 TTACTAGCTATGAGCTTCCAGGCCACTCTTCCATGGGCTATTTGTACGCTTAGAGTGGGA 627  
 Db 514 ATATTTATTTGATGATTCATGCAACCATTTGCCCTGGTCCACATGGCGGTGGAAATGGGG 573  
 QY 628 GA---ACTGGGTACCCCTCAGATC----- 647  
 Db 574 TATCCACTGATTAATCACTACGCCCGGATGCGAGTAATTTGTCTCACTGGAAGCAATGA 633  
 QY 648 -CAACACTGTGCTGATGATCAACATCACCACCA-----TGGTACCGAGAGTGTCTCA 699  
 Db 634 TCACGCGCCCGCAAAATTTACAAATGAAATTCACAAAGATCGGCTATCACAGAGTTGGGA 693  
 QY 700 ACTCTACTTTTGGAAACAGTTCTCCAAACAGGATGAATTTAGAGAGTCTCGGTGC 759  
 Db 694 ATGGTACTTTGTTAAGGAAGTGTGACAGAGAAACCAAAATATCGAGGAAGGAATCGGCT 753  
 QY 760 CCCCATCTGATCTGGTGTGTGTCTATTCATGATGATGATGATGATGATGATGATGATGAT 819  
 Db 754 ACCAAAGTGGAGTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 813  
 QY 820 CGCCGAGAGAGTCAAGAGTTCCGGCAAGCGGCTTACTCTCGGCTCTTCCCATAGT 879  
 Db 814 TCGTGGCGAGTGAAGAGTTCCGGCAAGCGGCTTACTCTCGGCTCTTCCCATAGT 873  
 QY 880 TGTCTATGATCACTTTATTCATCACCACCAATATCTCTCGGCTCTTCCCATAGT 939  
 Db 874 CATCATGGGTGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 933  
 QY 940 GTTCTGCTCAAGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 999  
 Db 934 CTACTTTTAAAGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 993  
 QY 1000 CACGCAAGTGTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1059  
 Db 994 AACTCAGTGTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1053  
 QY 1060 CAACGTTTCAAGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1119  
 Db 1054 CAACAGTGTGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1113  
 QY 1120 TACAGTGTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1179  
 Db 1114 GACCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1173  
 QY 1180 CAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1239  
 Db 1174 CGGAACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1233  
 QY 1240 CCCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1296  
 Db 1234 TCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1293  
 QY 1297 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1356  
 Db 1294 CATGCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1353  
 QY 1357 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1416  
 Db 1354 CATTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1413  
 QY 1417 CGGTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1476  
 Db 1414 CAGTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1473  
 QY 1477 AGATCACTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1536  
 Db 1474 GGACTTTTTCGCGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1533  
 QY 1537 GTTCTGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1596

Db 1534 TGGATGATTTATGACACAGATCGCTGTGCAAGAGCAATGATGATGATGATGATGATGATGAT 1593  
 QY 1597 GACTGGTGTCTACTGCGCTCTGCTGCGGCGGCTAATCACTGCTGATTAATGACAGCTGT 1656  
 Db 1594 AGTGGGCTGTACTGAGGCTCTGCTGAGCATCATTAACCCGCTGATATACATGTTAT 1653  
 QY 1657 GTTCTTCTACGCTCTTCTGCGCTCTAACAACCTGTGTTGGAGACACTACGATATACC 1716  
 Db 1654 TGTGATCTA---TTCTATGAGCGACTACCAACCTCAATCAACAAATATGCTATGCC 1710  
 QY 1717 GACTGCTGTATGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1776  
 Db 1711 AAATTTGCTTCTACTCAATTTGCTGCTAATTAACGCGCTTGTGATTTCTTACGCTGCTCAT 1770  
 QY 1777 TGGATGATTTCTTCTGATCAATTAACGCTGACCGAATCTTACGAGACATCAAGAA 1836  
 Db 1771 TTGATGATGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1830  
 QY 1837 AGCCTTCCACTCCAAACCTCATGAGGCTCCCG 1869  
 Db 1831 AGCCTTACCGCGAAAAAGAAATTTGGGAGACCCAG 1863

RESULT 10  
 AF369383  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Aedes aegypti.  
 Aedes aegypti.  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.  
 1 (bases 1 to 1996)  
 Dasher, M.K., Kohn, A.B., Harvey, W.R. and Stevens, B.R.  
 AEMAR, A Novel Amino Acid Transporter from Larval Aedes aegypti  
 Midgut  
 Unpublished  
 2 (bases 1 to 1996)  
 Dasher, M.K., Kohn, A.B. and Harvey, W.R.  
 Direct Submission  
 Submitted (09-APR-2001) Physiology, Whitney Laboratory, University  
 of Florida, 9505 Ocean Shore Blvd., St. Augustine, FL 32080, USA  
 Location/Qualifiers  
 1..1996  
 /organism="Aedes aegypti"  
 /db\_xref="taxon:7159"  
 1..1848  
 /note="NAAT"  
 /codon\_start=1  
 /product="amino acid transporter"  
 /protein\_id="AA013400.1"  
 /db\_xref="GI:20152861"  
 /translation="MATVSYSSMHPSPKPTKIALSGRTSNEFLCHVFOCPMSKRCGR  
 EPTFAENGGGAPLIPYIVLFAVIGRPVLYEMALGGTSSSVKIEWEPLFEGISIT  
 GDLVETTSVYSYSLALTIHYIFASASLPRATCKDNADNCVSSLYVNMEDS  
 NSTSGQKSSQIYFLDLVLEKESIDGIGAPKMLTLWILLANVYIFLYVGVCS  
 SGKAVFLAIPPYVLLIILRALTEADVGVEFTKPKWELLNFRVARSATTOLE  
 FSLVGMSSIIIMFSYNNFHNIYRDAMIYVTDFTSLGGMTFSLGLANLHDI  
 EDISVYSGTGLAFISYDIAKEDIVPOLPSELYFFMLFVLGVSAVALHGAITIA  
 FMDAPFKRYWHLALISTIGEGGLVYIHLVANGYSIMSTTEAFSDLRALITEMVA  
 IFWYIGLDMKNDIEPMOVRVGLYVAMAGSNNAFPDCHREHFLIVEYKPTVSGGO  
 YPEALIGCTVFIFGILQVILMAWYITGSTRSESWGKTSKAKPSOSOMGPECTEH  
 RKAWLKYEAKSRDDILITAKNHSIYKLCVLEPMIDETIQDLSRL"  
 BASE COUNT  
 ORIGIN  
 539 a 388 c 458 g 610 t 1 others

Query Match 12.8%; Score 366.8; DB 3; Length 1996;  
 Best Local Similarity 54.5%; Pred. No. 2.4e-69;  
 Matches 936; Conservative 1; Mismatches 753; Indels 28; Gaps 9;  
 QY 239 GAACGCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 298



Db 45 GCAGCCCAAAATGCGCCCTGAGTGTGCAACAGCTGATTTCTTATGTCATGATTTTCA 104  
 Oy 299 ACATCCGTTGGGTTGGGTACGTCGTGGCGTTCCCTTCATCGCGTACGAGATGAGAGA 358  
 Db 105 ATGTCCGTGTCTAGG--AAATGTGGCGTTTCCCTTACCGCGATGAATATGGGGGT 162  
 Oy 359 GGTGCTTCCCTGGTCCATACGTCATCTTCTTACTGTGGGCAAGCGTGTACTAC 418  
 Db 163 GGAGCGTTTATTTATTCCTACCCCATCGTCGTGTGTATAGAGACGCCGCTTATATATAC 222  
 Oy 419 TTAGATGTGTCTCGGACATTCAGTTCAAGAACTCTGTTAAAGTTGTGCAATTTCA 478  
 Db 223 CTGGAAATGTCTGGAGCAATTTACAGACAGACTCGTGAAATTTGGAGGTTGCG 282  
 Oy 479 CCGGCATGAAGATGATACGTCAGTACGTCGCGGCTGGGTTACATCTGCTCTAC 538  
 Db 283 CCACGTTCAAAAGGATCGGCAATTGGGCACTGTGGCGACTACTTCTGTGTAGCTAC 342  
 Oy 539 TAGCTGTGATCTGTGCTCTGTCTGATTTACTAGTATGAGCTTCAGGCGACCTCT 598  
 Db 343 TAGCTGTGATCTGTGCTCTGTCTGATTTACTAGTATGAGCTTCAGGCGACCTCT 402  
 Oy 599 CCATGGGCTATTTGTCAGCTCTAGTGG--GAGAACTGCTACCTCTAGATCCAACTT 655  
 Db 403 CCGTGGGCGACTGTGAAGATTAATGGGCTGATTAATGTGTAGACTCATCTGCTGTC 462  
 Oy 656 GCTGATCAGTCAACATCACA-----ATGTACAGCAGTGTCACTCTAC 706  
 Db 463 AATGAATGGGGATAGCAAGATTCATCCGACAGAAAGTGAAGAGCTTCTCAATATAT 522  
 Oy 707 TTTTGAAGACAGTCTCTCAACAAAGCGATGAATGAAGAGTCTCGTGGCCCTATC 766  
 Db 533 TTCTAGATATATAGTCTCAAGAAAGACAGATCGACGAGTATAGGGGCTCCGAT 582  
 Oy 767 TGTACTGTGTGTGTCTATTCATGCTAGTGGCTAGTGTGTGTGTGTGTGTGTGTGT 826  
 Db 583 TGAAGCTTATGCTTGT 642  
 Oy 827 GGAGTCAAGAGTTCGGGCAAGCGGCTACTCTCGGCTCTTCCCATAGTGTGTATG 886  
 Db 643 GCGGGAAGAGCTGTGAAGAGTACCTTCTTGGGATTTCCCATAGTGTGTGTGT 702  
 Oy 887 ATCACTTATTCATACACATCATCTCTGCGGCTGTACTGACGCGATCTGTCTTC 946  
 Db 703 ATCACTTATTCATTCGTCGCTTACCTGTGAAGACCGCTGACGAGATATCTTTTC 762  
 Oy 947 GTCAAGGCTCAATGGGCAAGCTCTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1006  
 Db 763 ATCAAGCAGAGTGGGGAATTTCTTATTCCTAAGTATGTGTGTGTGTGTGTGTGT 822  
 Oy 1007 GTGTCTTCTCTGACAGT 1066  
 Db 823 TTGTCTTCTCTGCTCTGT 882  
 Oy 1067 TTCAAGATATATCTAAGAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1126  
 Db 883 TTTCATACAAATATATAGATGATGATGATGATGATGATGATGATGATGATGATGAT 942  
 Oy 1127 TTCTGTCTGT 1186  
 Db 943 TTATTTAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1002  
 Oy 1187 GAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1246  
 Db 1003 GAAGACATTTCCAAAGGTGCAAAAGGTGCAAGGCTTGTAGCATTTATCATATTCGGAT 1062  
 Oy 1247 GCCATTTGCCAA--ACATTCACACCTCAGCTATCTGCTGTGTGTGTGTGTGTGTGT 1303  
 Db 1063 GCCATGCTTAATTTGACATTTGACACAGCTTTTCCGTGTATTTCTTCATGTTG 1122  
 Oy 1304 TCGTCTGT 1363

Db 1123 TTGTGTGGGCTAGGCTCGGAGTACCTTACATGAGCATATATATGATTTTGG 1182  
 Oy 1364 GACGGCTTCCCAACGCTACCCACGCTACATGTCAGAGATGAGACTGTTCTGGGTTTC 1423  
 Db 1183 GATGCTTCCCAAGGCAAAATATTTGGATCTGGGCTATTTATTTATCCAGATTTGAT 1241  
 Oy 1424 CTGCTTGGACTTGTTTACTGACACCGGCTGGACAAATATTTCTGTAGCTGTAGATCAC 1483  
 Db 1242 CTGCACAGATTTGGTGTATATACCTGTGTGGCCAAATGATATCTGATCTGTGTGTGT 1301  
 Oy 1484 TACGTTGACATCTTGTGCTTTTCTGGCCATTTCTGACCTGACAGCGCTGTGTGTG 1543  
 Db 1302 TACGAGG--CTTTTCTGATCTAGCTTACGTCACCTGATTAATGATTTGCAATCTTTGG 1359  
 Oy 1544 ATTTATGATTTGGAATCTGTGCTTACATGATTTAGTTCATGTTGGTAAAGACTGT 1603  
 Db 1360 ATTTATGATTTGGAATCTGTGCTTACATGATTTAGTTCATGTTGGTAAAGACTGT 1619  
 Oy 1604 GCTTAC--TGGGCTCTGCTGTGGGCTGATATCATCTCTGTATATAGACACTGTGTCTT 1662  
 Db 1420 CTCTACGTGGGCTATGTTGGGCTATATTAACGCCACTTTTCATGATTTGCCGTTTCAT 1479  
 Oy 1663 CTACGCTCTCTGCTCTTAAACAACCTGATGTGTGGAGACACTATACCCGACTGC 1722  
 Db 1480 CTATCTCTGTTGATATCAAAATGAGCCACATFACAGTGGCCAGAG--TATCCGGGCGA 1535  
 Oy 1723 TGTATATGTTCTGTGATCTGATGATTTTGGGCTAGTGTGTGTGTGTGTGTGTGTGT 1782  
 Db 1536 AGCTTATATTTGGGCTATTTGCTTTTATTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1595  
 Oy 1783 TGTATTTCTGTTTACAAATACGCTAC--CGAACCCTTACGAGACATCAAGAAGC 1839  
 Db 1596 TGTGTGACTGTGACAGAGATTTCTCAAAAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 1655  
 Oy 1840 CTTCACCTCAAAACCTCATGAGGCTCCGCTCGCCGAGAGACCTGATGAATGTATGCA 1899  
 Db 1656 GGCTAAGCCAAATTTCAACATGGGCTCCGTTATAGAAACACACGAAAGGCTGTAA 1715  
 Oy 1900 GTTCAAGGCTGAGAGCGAAAGCTTTAGACAAAGATGA 1937  
 Db 1716 ATCAAGAGAGAGCTTAATGAGGCGAGACGATATTA 1753

RESULT 11  
 AC015177  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 FEATURES  
 source  
 BASE COUNT  
 ORIGIN

AC015177 66809 bp DNA linear HTG 16-NOV-1999  
 Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\* in ordered  
 pieces.  
 AC015177.1 GI:6436158  
 HTG: HTGS PHASE2.  
 Drosophila melanogaster.  
 Drosophila melanogaster.  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidae; Drosophilidae; Drosophila.  
 1 (bases 1 to 66809)  
 Adams, M. and Venter, J.C.  
 Direct Submission  
 Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD, USA  
 This sequence was identified as CDM:10211923 by the submitter.  
 For further information on this sequence e-mail to fly@celera.com.  
 \* NOTE: This is a 'working draft' sequence.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

1. 66809  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 19034 a 14604 c 14832 g 18339 t





Best Local Similarity 56.2%; Pred. No. 3.2e-49;  
Matches 595; Conservative 0; Mismatches 409; Indels 55; Gaps 2;

```

QY 720 TTCTCCAAAGAGATGATGAGAGAGTCTCGGTGCCCCCATCTGACTTGGTGT 779
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7801 TTCTGCGGAGAGAGCGAGATGATGATGATGATGATGATGATGATGATGATG 7860
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 780 TGTGCTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 839
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7861 TGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7920
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 840 CCGGCAAGCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 899
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7921 CCGGCAAGCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7980
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 900 TCACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7981 TGGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8040
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 960 GGGGCAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1019
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8041 GGCACAAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1020 TGACAGTGTGACCGGACGATGATGATGATGATGATGATGATGATGATGATGAT 1079
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8101 TGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8160
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1080 TCTA-----CAGG 1087
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8161 TTTCACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8220
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1088 GATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1147
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8221 GATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8280
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1148 TTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1207
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8281 TTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8340
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1208 GCTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1264
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8341 AATGGAGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCG 8400
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1265 CAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1324
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8401 CTGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8460
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1325 TCGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1384
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8461 AATGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8520
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1385 ACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1444
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8521 AACTGAGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 8580
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1445 ACACCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1504
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8581 ACACCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8640
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1505 CTTTTCGCGCATTTTCTGAAGTCGAGGCGGTGTGTGAATTAAGATGGAATCTG 1564
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8641 CTGCTGTGCGCATTTCTGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 8700
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1565 TGGCTAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1624
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8701 TGGCTAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8760
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1625 GGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1684
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8761 GGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8820
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1685 AACCTGTGTGGAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1723
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 8821 CCCCTCAAGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8859

RESULT 13  
AC005894  
LOCUS

DEFINITION  
AC005894 196337 bp DNA linear INV 31-OCT-1998  
Drosophila melanogaster, chromosome 2R, region 46A1-46B2, P1 clones  
DS00050, DS00191, and DS05369, complete sequence.

AC005894 AC005435 AC005431 AC005441  
AC005894.1 GI:3818342

KEYWORDS  
SOURCE  
ORGANISM

Drosophila melanogaster.  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 196337)

REFERENCE  
AUTHORS

Celniker, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazek, R.G.,  
Butenoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,  
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,  
Hoskins, R.A., Houston, K.A., Hummel, S.R., Katta, K., Kearney, L.,  
Kim, E., Lee, B., Lewis, S., Li, P., Lomoth, M.A., Mazda, P.,  
Moshirefi, A.R., Moshirefi, M., Nixon, K., Pacle, J.M., Park, S.,  
Pfeiffer, B., Poon, L., Poon, E., Sequeira, A., Sethi, H., Smit, E.,  
Svirskas, R., Twomey, B., Wan, K.H., Weinburg, T., Zhang, R.,  
Zierman, L.L. and Rubin, G.M.  
Sequencing of Drosophila chromosome 2R, region 46A1-46B2  
Unpublished (1998)

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

2 (bases 1 to 196337)  
Celniker, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazek, R.G.,  
Butenoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,  
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,  
Hoskins, R.A., Houston, K.A., Hummel, S.R., Katta, K., Kearney, L.,  
Kim, E., Lee, B., Lewis, S., Li, P., Lomoth, M.A., Mazda, P.,  
Moshirefi, A.R., Moshirefi, M., Nixon, K., Pacle, J.M., Park, S.,  
Pfeiffer, B., Poon, L., Poon, E., Sequeira, A., Sethi, H., Smit, E.,  
Svirskas, R., Twomey, B., Wan, K.H., Weinburg, T., Zhang, R.,  
Zierman, L.L. and Rubin, G.M.  
Direct Submission  
Submitted (31-OCT-1998) Berkeley Drosophila genome project, MS  
64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,  
Berkeley, CA 94720, US

COMMENT  
JOURNAL  
REFERENCE  
AUTHORS

Sequence submitted by:  
Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory, MS 64-121  
Berkeley, CA 94720  
For further information about this sequence, including its location  
and relationship to other sequences, please visit our sequence  
archive web site (<http://www.fruitfly.org/sequence/>) or send email  
to [bdg@fruitfly.berkeley.edu](mailto:bdg@fruitfly.berkeley.edu).  
P1 library locations: 1-50, 2-95, 56-89.  
Location/Qualifiers  
1. 196337

FEATURES  
source

/organism="Drosophila melanogaster"  
/strain="Y2: cn bw sp"  
/db\_xref="taxon:7227"  
/chromosome="2R"  
/map="46A1-46B2"  
/clone="P1s DS00050 (D349), DS00191 (D345), and DS05369  
(D355)"  
/clone\_lib="P1 library, partial Sau3A in pAD10SACB11"  
/note="These P1s, DS00050 (D349) DS00191 (D345) DS05369  
(D355) were completed as a project. DS00050 extends from  
P1 end at bp 1 to P1 end at bp 79763. DS00191 extends  
from P1 end at bp 68,333 to P1 end at bp 154,851 DS05369  
extends from P1 end at bp 109,192 to P1 end at bp  
196,337."

BASE COUNT  
ORIGIN

56314 a 42176 c 42152 g 55695 t  
Query Match 9.6%; Score 274.6; DB 3; Length 196337;  
Best Local Similarity 56.2%; Pred. No. 3.2e-49;  
Matches 595; Conservative 0; Mismatches 409; Indels 55; Gaps 2;

QY	720	TTCTCCAAACAAACGATGGAATGGAAGAGAGCTCGGTGCCCCCATCTGGTACTTGGTGT	779
Db	187505	TTCTGGCGCAGAAAGCGAGCATGATGATGATGCAATATCCCAATGCCAGCTGGAGCTGGCGC	187564
QY	780	TGTGTTATTCATTCGATGCGATGCGATCATGATGTTCTCGAGATCGTCGCGCCGAGAGTCAAGATT	839
Db	187565	TGGCACTGTGCTGTGGCCTGGATGTCATGCGCGGGATTTATGTTCAAGGGGTCAAGAGATT	187624
QY	840	CCGGCAACCGGCTACTTCTCCCTGGCCTCTTCCCATACGTTGTGCATGATCATTATTC	899
Db	187625	CGGGCAAGGCTCTTCATCTTCCTGCCCCCTTTTCCCTACGCTGGAAGTGTGCTCTCTGG	187684
QY	900	TCACCAACAAATTCATCCCGCGGCTGACAGAGGATCCGTTCTGTCGACAGCCCAAT	959
Db	187685	TGCGGGCACTCAACCTTGGCGGGGCTTTGACGCGGCTGCTGTACTTCTCGCGCCCAAT	187744
QY	960	GGGGCAAACTCCTTGAGCTCGGTGTATGATGTTACTCAGCAGTCAACGAAATGTTCTCTC	1019
Db	187745	GGCACAAGTCTCTGGAAACCGCAGGCTCTGTAGCGCCGCTCCACCCAGGTTTCTCTCA	187804
QY	1020	TCACAGTGTGACCGGACCGATCATATGTTCTCTCTTCAACCGTTTCAGCATAAAT	1079
Db	187805	TGGCATCTGCTTTCGGGAACATCATATGATTTACGCTGTACAAACGCTTCGGCACAA	187864
QY	1080	TCCTA-----CAGG	1087
Db	187865	TTTACAGGTGATGAGATATCCGGTTTGAGTTCTTACGTGATTCCTGCTCCACAG	187924
QY	1088	GATGCTTGATTTGTTACGACCTTTGGACACCTTTACAAGTTTCTTGTCTGGGTGCACGATC	1147
Db	187925	GATGCCATATATGTATCAACCGCTGGACACTTTCACCTCCCTACTGTCCGCTGTGATATTC	187984
QY	1148	TTGCGTATCTTGTGATACCTGCGCTACGAACTCACTCAGAGTGGAGATGAGTGGCT	1207
Db	187985	TTTCGGATTTCTGGCAACCTGGCGTACGAGAACCAACACGACATCGCTAGTGGTTC	188044
QY	1208	GCTGGCGGTACCAAGCTTGTCTTCAATTTCATACCTGATGCCATTGCCAA---AACATTTC	1264
Db	188045	AATGAGGATCGGGCTGCTTTCATATCTTACCGGAGCCATTGCCAAATTTAAATGG	188104
QY	1265	CAACCTCAGTATTTCTGCGTGTCTTCTCGATGAGTGGTGGTGGTGTTCGCGCTCA	1324
Db	188105	CTGCCGAGCATATCTCCGTGTCTTCTTCCTCATGCTCTTGTCTTGGGCATTCGGCAGC	188164
QY	1325	TCCTGGCTGTCTATCGACTTTCACACATTGGGATGAGACGCGCTCCACAGTATACC	1384
Db	188165	AATGTCGGCATAGGCTCTCTGATGTCACCCGATCAAGGATCAATTTGGANACCTGAAG	188224
QY	1385	ACCGTCTCATATGTCAGCATGATCCTGTCTTCCGGTTTCTGTGACTGTATTACTGC	1444
Db	188225	AATGAGACTGTGGTGTGTTATGACCATTTGAGGCTACTTCTTGGGCTCTGTATATTC	188284
QY	1445	ACACCGGATGACATATATTTCTGAGCTTTAGATCACTACGCGTGAACATTTCTGTG	1504
Db	188285	ACACCCGGTGGCAGTTCTCTGCTCAACCTGCTCAGTACTTTGGCGTACAGTTCTGTGCC	188344
QY	1505	CTTTTTCGCGCATTTCTGAACTGCGAAGCGTGTCTGAGATTATGATGGAGATCTG	1564
Db	188345	CTGTGTCTGCGCATTTCTGAGCTGTGTACCATTTGCTGTGATCTACGGTGAAGGACTC	188404
QY	1565	TGCTTAGACATTTGATTTGATGGGTAAAAAAGACTGTGTACTGTGCGCTCTGCTGG	1624
Db	188405	TGCGGTACGATGAGTTTATGATTTGGCATCAAGACTTGTGCTGTATATGCAATCTGCTGG	188464
QY	1625	GGCGTAACTCTCTGCTATATATGACAGACTGTGTTCTTCAACGCTCTTCGCTCTAAC	1684
Db	188465	GCGGTCTCTCACTCTCTGCTATATGCTCAACCATTTCTCATCTACACCTGTTCTGTACAG	188524
QY	1685	AACCTGGTTCGGAGCACTAGATTAACCGACTGCT	1723
Db	188525	CCCTTCAATGATCAAGATTAACCTTACCAATTCGGGTGT	188563

RESULT 14	AE003832	242172 bp	DNA	linear	INV 28 -JUN-2002
LOCUS	AE003832				
DEFINITION	Drosophila melanogaster 2R section 18 of the complete arm,				
ACCESSION	Complete sequence.				
VERSION	AE003832 AE013599 AE002787				
KEYWORDS	HTG.				
SOURCE	AE003832.3 GI:21627589				
ORGANISM	fruit fly.				
REFERENCE	Drosophila melanogaster				
AUTHORS	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyrididae; Drosophilidae; Drosophila.				
	1 (bases 1 to 242172)				
	Adams,M.D., Celinker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,Y.H., Blazer,R.G., Champe,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J., Andrews-Planckhock,C., Baldwin,D., Bailey,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Benson,K.Y., Benois,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borrova,D., Borchan,M.R., Bouck,J.H., Brokstein,P., Brothier,P., Burtils,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,J.E., Evangelista,C.C., Ferraz,C., Ferreira,S., Fleischmann,W., Foster,C., Gabrielian,A.E., Garg,N.S., Gelbart,W.M., Glasser,K., Glodok,A., Gong,F., Gottrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Helman,T.J., Hernandez,J.B., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wel,M.H., Ijzerman,C., Jalali,M., Kalush,F., Karpén,G.H., Ke,Z., Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Krivitski,S., Kul,P.D., Lai,Z., Lascko,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Musken,D.R., Paclob,J.M., Palazolo,M., Piltman,G.S., Pan,S., Nussken,D.R., Paclob,J.M., Palazolo,M., Piltman,G.S., Pan,S., Pollard,J., Puri,V., Reese,M.G., Reinetti,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spadling,A.C., Stapleton,M., Strong,R., Sun,X., Svirskas,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.T., Wasserman,D.A., Weinstock,G.M., Weissbach,J., Williams,S.M., Woodage,T., Wolley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yen,R.F., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O., Gilbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.				
TITLE	The genome sequence of Drosophila melanogaster				
JOURNAL	20196006				
PMID	20196006				
REFERENCE	10731132				
AUTHORS	2 (bases 1 to 242172)				
	Amanatides,P.G., Brandon,R.C., Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A., Carlson,J.K., Center,A., Champe,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Ferritera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A., Gonzalez,M., Houck,J.J., Hoskins,R.A., Hostin,D., Howland,T.J., Idegawa,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Paclob,J., Parasas,V., Park,S., Patel,S., Pfeiffer,B., Phouanavong,S., Piltman,G.S., Puri,V., Richards,S., Scheeler,F., Svirskas,R., Strong,R., Svirskas,R., Tector,C., Tyler,D., Williams,S.M., Zaveri,J.S., Smith,H.O., Venter,J.C. and Rubin,G.M.				
TITLE	Sequencing of Drosophila melanogaster genome				
JOURNAL	Unpublished				
REFERENCE	3 (bases 1 to 242172)				



QY	900	TCACCAACATTCCTCCGCCGCGCTGCTACGACGGACATCTGTTCTTCTGTCACGCCCTCAAT	959		
Db	191021	TGCGGGCACTCACTTGGTCCGGGGCGCTTCGACGGCGTGGTACTTCTGCGGCCCAAT	191088		
QY	960	GGGCGCAATCTCTTGAGCTCGGGGTATGGATGCTACACAGACGACGAAGTCTTCTCTC	1019		
Db	191081	GGCACAAGCTCTCGAACCAGCGAGTCTGGTACGCCGCCCTGACCCAGGTTCTTCTCAC	191140		
QY	1020	TCACAGTGTGACACCGGACCGATCATGTTCTCTTACCAACGGTTCTCAGACATATA	1079		
Db	191141	TGGCCATCTGCTTCGGGAACATCATATGATGTAGCCCTGTACAAACCGCTTGGCCACAAC	191200		
QY	1080	TCCTA-----CAAG	1087		
Db	191201	TTTACAGGTGAGTGGAGATATCCGGTGGAGTCTCTTACCTGATTCCTGCTCCACAG	191260		
QY	1088	GATGCTTGGATTTGTTACGACTTTTGACACTTTTACAAATTTCTGCTGGGTCCAGATC	1147		
Db	191261	GATGCCAATATGCTGACCAACGGTGGACACTTTCACCTCCCTACTGTCGGTGTGATTAAC	191320		
QY	1148	TTCCGATTCCTTGGTAACTCCGCTGACACTCAACTCAAGAGTGGAGATGTGTCGGT	1207		
Db	191321	TTCCGGAATCTGGGAACTCTGGCGACGAAACACACACCGACATCTCCTAGTGTGGTC	191380		
QY	1208	GCTGGCGGACAGCTTCTCTTCACTTTATATACCGGATGCCATTTGCCAA---AAATTC	1264		
Db	191381	AATGGAGTCCGGGCCCTGGCTTTCATATTTACCCGGAGCCAAATTCCAAATTAAGTG	191440		
QY	1265	CACCTCAAGTATTCCTGGTGTCTTCTCTGATGATGCTGGTGGGTATCGGCTCA	1324		
Db	191441	CTGCCGCAAGCTATTCCTCCGCTGTCTTCTCTCATGCTCTTGCTCTGGGCAATGGCAGC	191500		
QY	1325	TCCGGGCTCTGCTATTCACACTTTCACAACATTTGGCGATGGACCGCTTCCACGTGATCC	1384		
Db	191501	AATGTGGGCAATGGCCTCTCTGCAATGTCACACCGATCAAGATCAATGTCGACACCTGAAG	191560		
QY	1385	ACCGCTCATATGCTACAGCATGACACTGATCTTCTCGGTTCTGCTGGATTTGTTACTG	1444		
Db	191561	AACGTGACTGTGGTGGTATTAACCATTTGTAAGCTACTTCTGGGCTTCTGTATATC	191620		
QY	1445	ACACCGGTGGCAATATATTTCTTGAAGCTTGTATATCACTACGATGGGAATCTCTTGTG	1504		
Db	191621	ACACCGGTGGCAATCTCTGCTCAACCTGTGCTACTTGTGGGCTGACGTTGCGGCC	191680		
QY	1505	CTTTCTCGCCATTTCTGAATCTGCACAGGCGTCTGATTTATGATTTAGATTTGAGATATCG	1564		
Db	191681	CTGTGCTGGCCATTTGAGAGCTGTGACCATTTGCTGATCTAGGCTGTGAAGGACATC	191740		
QY	1565	TGCTTAGACATTTGACTCATGTTGGGTAAAAAGACTGTGGTCTTACTGCGCTGTCTGCTG	1624		
Db	191741	TGCCGTGACGTGAGATTCATGATTTGGACATCAAGACTTGGCTACTATACGATCTGCTG	191800		
QY	1625	GGCGAATACATCTCCGCTATATATACAGACTGTGTTCTTACAGCTTCTCGCTCTAAC	1684		
Db	191801	GCGGCTGCTACTCTCTCATATGTCACATTTCACTCATCACTTGTTCTGTACAG	191860		
QY	1685	AACCTGTGTTCGAGACACTACATACCGCATCTGCT	1723		
Db	191861	CCCCGACAGTACAAAGATTTACACCTACCAATCGGCTTT	191899		
RESULT 15	AC018319	18567 bp	DNA	linear	HTG 09-DEC-1999
LOCUS	AC018319/c				
DEFINITION	Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered pieces				
ACCESSION	AC018319				
VERSION	AC018319.1				GI:5552872
KEYWORDS	HTG; HTGS; PHASE2.				
SOURCE	Drosophila melanogaster.				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	source
Neoptera: Emdopterygota: Diptera: Brachycera: Muscomorpha: Ephydroidea: Drosophilidae: Drosophila.	1. (bases 1 to 18567)	Adams, M. and Venter, J.C.	Direct Submission	Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA	This sequence was identified as CDM:10214336 by the submitter. For more information on this record e-mail to fly@celera.com. * NOTE: This is a 'working draft' sequence. * This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved.	Location/Qualifiers
						1. 18567
						/organism="Drosophila melanogaster"
						/db_xref="taxon:7227"
BASE COUNT	5705 a	3897 c	3763 g	5202 t		
ORIGIN						
Query Match	9.4%	Score 270.2:	DB 2:	Length 18567:		
Best Local Similarity	58.0%:	Pred. No. 2.8e-48:				
Matches	498:	Conservative	0:	Mismatches 358:	Indels	3: Gaps 1:
Db	636	CTCAACTACTATTTTGGAGAACAGTTCACCAACAAAGGATGGAATTAAGAGAGTTCG	755			
Db	6023	CGCACTACTACTTGGACAGCATCGTGTGAATGAAGAGATTCCCTGGAGAGGACATCG	5966			
Db	756	GTGCCCCCATCTGCTACTTGGTGTGTGTCTATTATGATGCGTCACTGTTGGAG	815			
Db	5963	GATTAACCAAGTGTGATTGGCTTAATGCTGGGCATTTTCGTGCTAACCGTAACTTAA	5904			
Db	816	TGCTGCCCGGAGAGTCAAGAGTTCGGGCAAGCGGCTACTTCCCTCGCGCTTCCCAT	875			
Db	5903	TGATTAATTCGGGGAATGAAGATTGAGAAAGCGGCTATGCTCGCTTTTCCGT	5844			
Db	876	ACGTTGATGATCACTTTATTATCATCACACATCATCTCCCGGCTCTACTGACGCA	935			
Db	5843	ACGTTGATGATTTTATCCCTGCTCGAGAGACATCATTAACAGAGTCTTACAGCGAG	5784			
Db	936	TGCTTTTGTGTCACGCTCAATGGGGGAAACCTTGACTCGGTATGATGATCAG	995			
Db	5783	TTATGTACTTTCTGACTCACAGTGGGAAACTTTTGGACCGCAGGTGGTACATG	5724			
Db	996	CAGTCACGAAGTGTCTTCTCTGACAGTGTGACCGGACGATCATATGTTTCCCT	1055			
Db	5723	CCGTACCCAGATGTTTCTTCCCGGCGCTGTGTTGGTGTATCATATGATCTCT	5664			
Db	1056	CTTCAACAGGTTTCAGATATATCTACAGGATGCTTGATTTGTAACACTTTGGACA	1115			
Db	5663	CGTCAACCGTTTGGCTACAAATGCTATCGGGATGCCAATATATGTAACCAACCTCGACA	5604			
Db	1116	CTTTTACAGTTCCTTGTCTGGGTCACAGATCTTCTCGATCTCTGGTATCCCTCGATG	1175			
Db	5603	CATTACACTACTGCTCTCGGGGTGATATCTTTGGAAATCTTGGAAACCTGGCTACG	5544			
Db	1176	AACCTACCTCAGAGTGGAGATGTTGCTGCTGGCGGTACCAAGTCTTGCTTCAATT	1235			
Db	5543	AATGGGACCAACAGACATTCGCCAGTGTGTTAAGGAGAGCTCTGGATTTGGCTTACCT	5484			
Db	1236	CATACCCGATGATCCCAAA--ACATTTCAACCTCAGCTTTCCTCGGTGCTGCTCT	1232			
Db	5483	CTTATCCGGATGCCATTGTCACAAAGTTCAAAATGTTTTCCTCAAGTCTTTTCTGCTCTCT	5424			
Db	1293	TCTGATGATGCTGGGTATTCGGCTCATCCGATCTTCTGATTCGACTTTCAACA	1352			
Db	5423	TCCGATGCTTTCATGTTGGCGTGTGTATCAATGTTGGGATGATGATGATGATGATG	5364			
Db	1353	CATGGGATGAGAGCGCTTCCAGCTGATCCACCGCTCTACATGTCAGCATGACTGTT	1412			
Db	5363	CCGTGCTCAAGATCATGCTGTGAACCTCAAGCTGTGGATTAATATGATGATGATGATG	5304			
Db	1413	CTTCCGATTCCTGCTTGGACTTTTACTGACACCGGGGTGACATTAATTTCTTAGC	1472			

Db 5303 TGATAGGATTCTGCTGGTGGGATTGATTACATCACTCCCGAGGTCACACATATATCATCAGC 5244  
QY 1473 TTGTAGATCACTACGGGTGAACATTCCTTGCTTTCTGCGCCATTCTGAACTCGACG 1532  
Db 5243 TTATGACCTTTCACGGCGCTCACCCTTGCTCCCTGGTATCGGCCCATCTTGAGAGTCATG 5184  
QY 1533 GCGTGTCTGATTTATG 1551  
Db 5183 CCGTGGGCTGGATCTACGG 5165

Search completed: July 26, 2003, 02:02:05  
Job time : 7286 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2003, 22:31:53 ; Search time 621 Seconds  
(without alignments)  
103771.522 Million cell updates/sec

Title: US-09-991-458-1

Perfect score: 2860

Sequence: 1 ggcacgaggttactgttg.....ataaaaaaaaaaaaaa 2860

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

N\_Geneseq\_101002:\*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*

2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*

3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*

4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*

5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*

6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*

7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*

8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*

9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*

10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*

11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*

12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*

13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*

14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*

15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*

16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*

17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*

18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*

19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*

20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*

21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*

22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*

23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*

24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	455.6	15.9	1956	ABL03173	Drosophila melanog
2	395.4	13.8	2036	ABL14421	Drosophila melanog
3	393.8	13.8	2205	AAD03450	Drosophila melanog
4	391.8	13.7	2473	AAD09681	Drosophila melanog
5	391	13.7	2013	ABL05887	Drosophila melanog
6	383.4	13.4	1992	ABL07751	Drosophila melanog
7	383.4	13.4	2030	AAD09683	Drosophila melanog
8	378.6	13.2	2062	ABL28723	Drosophila melanog
9	378.6	13.2	2108	AAD09684	Drosophila melanog

c	10	274.6	9.6	4540	23	ABL03172	Drosophila melanog
c	11	270.2	9.4	4385	23	ABL07750	Drosophila melanog
c	12	247.4	8.7	5270	23	ABL14420	Drosophila melanog
c	13	231.4	8.1	4301	23	ABL05886	Drosophila melanog
c	14	229.2	8.0	6148	23	ABL28722	Drosophila melanog
c	15	222	7.8	2817	19	AAV01590	Rat glycine transp
c	16	211	7.4	1983	13	AAQ28118	Human norepinephr
c	17	209.6	7.3	2728	18	AAV58853	cDNA encoding high
c	18	209.6	7.3	2728	19	AAV28861	Rat proline transp
c	19	209.4	7.3	1854	22	AAH28082	Nucleotide sequenc
c	20	209.4	7.3	1854	22	AAH28086	Nucleotide sequenc
c	21	209.4	7.3	3404	13	AAQ26683	Rat dopamine trans
c	22	207.8	7.3	1854	22	AAH28083	DNA encoding human
c	23	207.8	7.3	1854	22	AAH28087	DNA encoding human
c	24	207.8	7.3	1985	14	AAQ41172	Dopamine transport
c	25	204.8	7.2	3919	14	AAQ53413	Sequence of the p
c	26	204.8	7.2	3946	24	ABL68342	Sequence of the p
c	27	201.4	7.0	2640	13	AAQ26664	Rat 5HT transp
c	28	197.8	6.9	2278	14	AAQ41055	Human glycine tran
c	29	196.4	6.9	2394	21	AAQ61440	Human glycine tran
c	30	196.2	6.9	2278	16	AAQ94016	DNA encoding a hum
c	31	196.2	6.9	2415	16	AAQ94017	Human glycine tran
c	32	194.8	6.8	2394	21	AAQ61439	Human glycine tran
c	33	194.8	6.8	2394	21	AAA91842	Human glycine tran
c	34	194.8	6.8	2394	21	AAA92001	Human glycine tran
c	35	194.8	6.8	2397	19	AAV22932	DNA encoding a hum
c	36	193.2	6.8	2394	21	AAA61441	Human glycine tran
c	37	193.2	6.8	2397	19	AAV22915	Human GLYT-2 trans
c	38	193.2	6.8	2397	19	AAV22917	Human GLYT-2 trans
c	39	193.2	6.8	2397	19	AAV22920	Human GLYT-2 trans
c	40	193.2	6.8	2397	19	AAV22921	Human GLYT-2 trans
c	41	193.2	6.8	2397	19	AAV22922	Human GLYT-2 trans
c	42	193.2	6.8	2397	19	AAV22923	Human GLYT-2 trans
c	43	193.2	6.8	2397	19	AAV22924	Human GLYT-2 trans
c	44	193.2	6.8	2397	19	AAV22925	Human GLYT-2 trans
c	45	193.2	6.8	2397	19	AAV22926	Human GLYT-2 trans

## ALIGNMENTS

RESULT 1	ABL03173	standard; cDNA; 1956 BP.
ID	ABL03173	
AC	ABL03173;	
DT	26-MAR-2002	(first entry)
DE	Drosophila melanogaster expressed, polynucleotide seq ID NO 4001.	
XX	Drosophila: developmental biology; cell signalling; insecticide;	
KW	pharmaceutical; gene; ss.	
KW	Drosophila melanogaster.	
OS	Drosophila melanogaster.	
XX	WO200171042-A2.	
PN	27-SEP-2001.	
XX	23-MAR-2001; 2001WO-US09231.	
PF	23-MAR-2000; 2000US-191637P.	
XX	11-JUL-2000; 2000US-0614150.	
PR	(PEKE ) PE CORP NY.	
XX	Venter JC, Adams M, Li PMD, Myers EW;	
PI	WPI: 2001-656860/75.	
XX	P-PSDB; ABB59070.	
DR	New isolated nucleic acid detection reagent for detecting 1000 or more	
XX		
PT		







AC AAD03450;  
 XX 13-JUN-2001 (first entry)  
 XX Drosophila melanogaster bioamine transporter 1 (BTL) homologue cDNA.  
 DE  
 XX Fruitfly; transporter protein; bioamine transporter 1; BTL; screening;  
 KM therapy; molecular transport disorder; biopesticide;  
 KM Invertebrate model; ss.  
 XX Drosophila melanogaster.  
 OS  
 FH Key Location/Qualifiers  
 FT CDS 159..2048  
 FT /\*tag= a  
 FT /product= "Fruitfly bioamine transporter 1 homologue"  
 XX MO200118178-A1.  
 XX 15-MAR-2001.  
 XX 08-SEP-2000; 2000WO-US24598.  
 XX 09-SEP-1999; 99US-0153461.  
 XX (GENO-) GENOPTERA LLC.  
 PA  
 PI Ebdens AJ, Keegan KP, Winslow JW;  
 XX WPI: 2001-235196/24.  
 DR P-PSDB: AAE00297.  
 XX  
 PT Drosophila melanogaster Bioamine Transporter 1 (BTL) nucleic acid and  
 PT protein, useful in screening assays to identify candidate compounds  
 PT which are potential pesticide agents or therapeutics that interact with  
 PT BTL proteins -  
 XX  
 XX Claim 2; Page 48-49; 53pp; English.  
 XX  
 CC The present sequence is a Drosophila melanogaster transporter protein  
 CC homologue, bioamine transporter 1 (BTL). The BTL is used in screening  
 CC assays to identify candidate compounds which are potential pesticide  
 CC agents or therapeutics that interact with BTL proteins. It can also be  
 CC used to genetically modify metazoan invertebrate animals resulting in  
 CC BTL expression or mis-expression. It can also be used in methods for  
 CC identifying new drug targets, therapeutic agents, diagnostics and  
 CC prognostics useful in the treatment of disorders associated with  
 CC molecular transport across the membrane. The BTL nucleic acid or its  
 CC fragments, such as an antisense sequence or double stranded RNA, may be  
 CC used as a biopesticide to inhibit BTL function. The invertebrate model  
 CC organisms such as Drosophila melanogaster are useful for rapidly carrying  
 CC out large-scale systematic genetic screens and therefore for analysing  
 CC expression and mis-expression of BTL protein.  
 CC  
 XX  
 XX Sequence 2205 BP: 478 A; 517 C; 570 G; 640 T; 0 other;  
 SQ  
 Query Match 13.8%; Score 393.8; DB 22; Length 2205;  
 Best Local Similarity 54.1%; Pred. No. 1.7e-76;  
 Matches 872; Conservative 0; Mismatches 732; Indels 9; Gaps 3;

Db 516 TTCTCGAGCGGTGTTGATTCAGAGCCCTCGAATGATCATATATAGGGGTATTTGCC 575  
 Qy 500 TACGTCACAGTCCGCGCTGCTGATATCTCTTACTAGTGGTATCTGTGCTC 559  
 Db 576 TATGTCAGGTGATTCACCGCTTGGCCACCACCTACCTGCTGATATAGGCTTG 635  
 Qy 560 TGTCTATTTACTTACTATGACCTTCAGGGCAGCTTCATGGGCTATTTGTGACGCT 619  
 Db 636 ACCATGAGATTTTGGCTGCTATGAGTGAAGTCTGCTGACATATTTGCTCTG 695  
 Qy 620 GAGTGGAGAACATGCTTACCTCAGATCCAAACATTTGCTCATCAACATCAACAC 679  
 Db 696 GATGGGGCAAGATTTGGCTGGAGGAGGAGCCAGAGCTGCCAATGATCATCATTTG 755  
 Qy 680 AATGTACACAGAGTGTCTCACTTCTTTTGAAGACAGTTCTCCAAACAGCTGAG 739  
 Db 756 CAGGGGGTTTCCCTCCCGAGCTGTTTTCACCAACAGTTCTGAGGGAGCCGAATCG 815  
 Qy 740 ATTGAAG---GAGGTCCGGTGGCCCATCTGCTACTGCTGTTGTTGATTCATCGCA 796  
 Db 816 CTGAGAGATATATGGCTGGGTACTCCAGTTGGGATCTGTTGTTGTTCTCTGGCCACT 875  
 Qy 797 TGGCTATGCTGTTGGAGTCTGCGCCGAGAGTCAAGAGTCCGGAAGCGGCTTAC 856  
 Db 876 TGGGTATCATTCGACAGATTTTGTCAAGAGCATTCGCGAGTTGGAAAGCTTCTAT 935  
 Qy 857 TTCTCGCGCTCTCCCATACGTTGATGATATCATTTATTCATCAACCAATATCTCG 916  
 Db 936 TTCTCGCGCTCTCCCATACGTTGATGATATCATTTATTCATCAACCAATATCTCG 995  
 Qy 917 CCGGTCCTACTACGAGCATCTGTTCTGTCAGCCCTCAATGGGCGAAATCTTTGAG 976  
 Db 996 CAGAGTCCCTGGCAGGAATTTGTACTTCTTAAGCCCAATGTGCCAGCTGTCAAT 1055  
 Qy 977 CTGCTGATGATGATCAGCAGTCAAGTCTTCTCTGACAGTGTCAACCGGA 1036  
 Db 1056 CCCACAGTGTGATCAGGGGCGCATCACAGATTTCTTCGCTGGCCATCTCTCGGA 1115  
 Qy 1037 CCGATCATGATTTCTCTCTTCAACGGTTGATGATATTTCTAATGATGATGATG 1096  
 Db 1116 ACGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1175  
 Qy 1097 ATTGTACGATTTGAGACCTTACAGTTTCTGTGTGGGTGACAGATCTTGGTATC 1156  
 Db 1176 ATATATCCACTATAGATCTGCTTACATTCATCTGCTGATGATGATGATGATG 1235  
 Qy 1157 CTGTGATACCTGCTGATCAGATCTCACTGAGAGTGGAGATGTGCTGCTGGCGGT 1216  
 Db 1236 CTGGGCAACCTGCTGATGAGACCAACCAAGATATTTGCGAGGTGTCAAAGTGA 1295  
 Qy 1217 ACCAGTTGCTTATTTATATCCCTGATGATGATGATGATGATGATGATGATGATG 1273  
 Db 1296 GCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1355  
 Qy 1274 CTATTCGCTGCTGCTTCTCTGATGATGATGATGATGATGATGATGATGATGATG 1333  
 Db 1356 CTGTGCGCTGCTGCTTCTCTGATGATGATGATGATGATGATGATGATGATGATG 1415  
 Qy 1334 CTGTATGACTTTCAACATATGAGATGAGAGCGCTTCCAGCTGATACCATGCTTAC 1393  
 Db 1416 ATGGCTTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1475  
 Qy 1394 ATGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1453  
 Db 1476 CTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1535  
 Qy 1454 GACAAATATATTTTGAAGCTTGAATGATGATGATGATGATGATGATGATGATGATG 1513  
 Db 1536 GAGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1595  
 Qy 1514 GCCATTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1573



QY 875 TACGTGTGATGATCACTTATTATTCACCAACATCATCTGCCGTCTACTGACGGC 934  
 DB 1047 TACGTGTGATGATCACTTATTATTCACCAACATCATCTGCCGTCTACTGACGGC 1106  
 QY 935 ATCTGTCTCTCTGACGCTCTCAATGAGGGAACCTCTTGTAGTGTGATGATGATCTCA 994  
 DB 1107 ATCTGTCTCTCTGACGCTCTCAATGAGGGAACCTCTTGTAGTGTGATGATGATCTCA 1166  
 QY 995 GCAGTACGCAAGTGTCTCTCTGACAGTGTGACCGGACCGATCATCATGTTCTCC 1054  
 DB 1167 GCAGTGTGAGAGCTCTCTCTGACAGTGTGACCGGACCGATCATCATGTTCTCC 1226  
 QY 1055 TCTTACACGCTTTCACACATATATCTACAGGATGCTGTGATGTTAGACTTGTGAC 1114  
 DB 1227 TCTGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1286  
 QY 1115 ACCTTACAGATTTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1174  
 DB 1287 ACCTGTACAGATTTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1346  
 QY 1175 GAATCTACATGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1234  
 DB 1347 GAATCTACATGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1406  
 QY 1235 TCATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1291  
 DB 1407 TCATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1466  
 QY 1292 TTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1351  
 DB 1467 TTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1526  
 QY 1352 ACATGCGGATGAGAGGCTTCCACGCTGTACACCGCTGTACATGATGATGATGATGAT 1411  
 DB 1527 ACATGCGGATGAGAGGCTTCCACGCTGTACACCGCTGTACATGATGATGATGATGAT 1586  
 QY 1412 TCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1471  
 DB 1587 TCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1646  
 QY 1472 CTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1531  
 DB 1647 CTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1706  
 QY 1532 GCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1591  
 DB 1707 GCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1766  
 QY 1592 AAAAAGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1651  
 DB 1767 GCGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1826  
 QY 1652 ACTGTGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1711  
 DB 1827 ATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1883  
 QY 1712 TACCGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1771  
 DB 1884 TACCGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1943  
 QY 1772 CCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1831  
 DB 1944 CCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2003  
 QY 1832 AAGAAAGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1891  
 DB 2004 AAGAAAGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2063  
 QY 1892 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1938  
 DB 2064 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2110

RESULT 5  
 ABL05887  
 ID ABL05887 standard; cDNA; 2013 BP.  
 XX  
 AC ABL05887;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 12143.  
 DE Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PE 23-MAR-2001; 2001MO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 PI Venter JC, Adams M, Li PWD, Myers EM;  
 XX  
 DR WPI; 2001-65860/75.  
 DR P-PSDB; ABB61784.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 PS  
 Claim 1; SEQ ID NO 12143; 21bp + Sequence Listing; English.  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
 CC sequences (AB101840-AB116175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
 CC  
 SQ Sequence 2013 BP; 414 A; 511 C; 553 G; 535 T; 0 other;  
 XX  
 Query Match 13.7%; Score 391; DB 23; Length 2013;  
 Best Local Similarity 53.2%; Pred. No. 6.6e-76;  
 Matches 949; Conservative 0; Mismatches 765; Indels 69; Gaps 3;  
 QY 222 ACACGACTAGAGGCTGTAACCGCCAGACATGATGATGATGATGATGATGATGATGATGAT 281  
 DB 170 AACAGATGCGAGAACCGCCAGCGGAGGACACCAATTTGGGCAATGAGCTGTGAGTTTC 229  
 QY 282 TGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 341  
 DB 230 TGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 289  
 QY 342 CGTACAGAAATGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 401  
 DB 290 CCGTATGAGAAATGAGAGGAGGCGCTTCTTATACCTTACATGATGATGATGATGATGATGAT 349  
 QY 402 GCAAGCTGTGTACTACTTGAAGTGTCTCTGACAAATTCAGTTCAAGAACTCTGTGA 461  
 DB 350 GCAAGCCAGTACTATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 409  
 QY 462 AAGTTTGTCAATTTACACCGGCCATGAAAGTACTGATGATGATGATGATGATGATGATGATGAT 521

D	b		410	AGATTGGCGGTGGTGGCGGGATTTGGGGGGGTGGGCTATGAGCCAGGCTTGGCACC	469
O	y		522	GTTACATCTCTTACTACAGTGGTGATCTGTGCTCTGTCTGTAATTACTTAGTANGA	581
D	b		470	TCTGCATATCTCGACTATATCTTCGGCTGGGCGCTGACCTTACTATCTCTTGTGT	529
O	y		582	GCTTCAGGCGACCTTTCACGGGGCTATTGTGCACCGTAGGGGAGAACTCCGNACCT	641
D	b		530	CCTTCACATCGGAATGGCTGGTGTCTCTACTGTGCGAGATGGACCAATTCGTGA	589
O	y		642	C-----	642
D	b		590	CGAGCCGACAGATATGTGACAACTGTGCTTACGGGGGTATTCCTTGGCCAAATGAATGG	649
O	y		643	----AGATCCAAACACTTCTGCTGATCAGTCAACAAATATCCAAATGTACCAAGCTGCTC	698
D	b		650	CCAGAAATCTCAGCGGGCAATTTGGCCAAAGTGAACAGAGAAATCCAGAGACACTCCG	709
O	y		699	AACTTACTTTTGTGAAACAGTTTCCCAACAAGAGATGGAATTGAAGAGGTCCGGG	758
D	b		710	AACTTATTTCTGTAATGATGATCAAGAGAAACTGTGACATCTCGGACGGCTTGGTG	769
O	y		759	CCCCATCTGTACTTGGTGTGTGTCTATTTCATTCGATGECTCATGAGTGTTCGGAGTCG	818
D	b		770	ATCCCGATTTGAAGGTGACACTTGTGCCCTGTGCTGGCTCGCTGGGTGTCACTTTTCTG	829
O	y		819	TGCGCCGAGAGATGACAGAGTTCCGGCAAGGGGCGCTTCTTCCTGCGCTTCCCATACG	878
D	b		830	TCATGCGAGGGGTGAAGATTCGCGCAAGGGGCGCTTACTTCTGGCGCTGTCCCTACG	889
O	y		879	TTGTGATGATCACTTATTCATCAACACAAATCATCTCGCGGTCTACTGACGCAATCC	938
D	b		890	TGGTGCTCTTGCTCTGCTGTGATCCGGTGGCGCTTACCTTGAAGGAGACAGCGATGGCATTC	949
O	y		939	TGTTCTTGCTGACGCTCAATGAGGGGGAACCTCTGACTCGGCTATAGTACTACAGAG	998
D	b		950	TGTTCTTTCTTGAGGCGCACAGTGGGGTGAAGTCTTGAACCCACGCTTTGAAAGAGGCG	1005
O	y		999	TCACGCAAGTGTCTCTCTCTGACAGTGTGACCGGACCGCATCATGTTCTCTCT	1058
D	b		1010	TTGTGACATGCTTCTTCTCGCTGGGCGGGGCTCGGACCATTAATCATGTTCCGCTGT	1065
O	y		1059	ACAACGTTTCAGACATATATATCTACAGGATGCTTGATTTGTAACACTTTGGACACT	1118
D	b		1070	ACAATCGATTTGATCAGGAATCTATAGGATGTCATATGTAACCACTGACACAGC	1129
O	y		1119	TTACAAGTTCTTCTGCTGGTGACAGATCTTCGGATCTTGATTAACCTCGGTACGAGAC	1178
D	b		1130	TGACACGCTCTCGGAGAGTATMAAGATATTTGGCATTTGGGCATCTGGCGCAATTC	1188
O	y		1179	TCAACTAGAGGTGGAGATGTGTCGCTGGGCGGTACCAAGTCTTCTTCATTTAT	1238
D	b		1190	TGCGAATCGAAMAATCAGAGATGTGGTGCMAAGTGCGACGGGATTTGCTTCATCTGT	1249
O	y		1239	ACCGTGAATGCCATTGGCCAA--AAATTTCAACCTCACTATTCGCGGTGCTGTCTTCC	1299
D	b		1250	ATCCGGATGCAATATCAAAAGTTCCAGCGGGGTGCGGACACTTTTCTCGGTTCTGTCTCT	1305
O	y		1296	TGATGATGTGGTCTGGGTATCGGCTCATCCGTCCTCTATTCGACTTTCACACATAT	1355
D	b		1310	TCATGCTGTTGCTCTGGGCAATTTGGATCAATTTGGCCCTCGACAGCAACATTTGACACA	1365
O	y		1356	TGGGGATGGAGCGCTTCCACAGTGTACCCACCGCTCTACATGTCAAGCATGACCTGTTCTT	1415
D	b		1370	TCATTTGGATCACTATCAAGGATGGAAGTACTGGAAAGTGGCACTGACCACTCTGTGT	1429
O	y		1416	GCGGTTTCCTGCTTGGACTGTTTATCTACGACACCGGGGAGCAATATATTTCTGAGCTTG	1475
D	b		1430	GCGGTTCCTTATGAGGTTTGGTCTACGTTAACCGGGGGGACAGATGATCTTCACTTTGG	1489
O	y		1476	TAGATCACTACGGTGGAAATCTCTTGTGCTTTCTGCGGCATTTTGAAGTCCGAGGCG	1535

Db 1490 TGGACTTCTAATGCGGGAACCAATGCGTCTTCATCTCTGGCCATTTTGCACGTGCGTGGCA 1549

QY 1536 TGTTCGTGATTATGATTTGGAGATCTGTGCTAGACATGATGATTCATGTGGGTAAAA 1595

Db 1550 TCGTGTGGGTTTATGTCTCTGCAAAACTTTTGCATGACATGAGTTTATGTGCATGCCC 1609

QY 1596 AGACTGTGCTTACTACGCGCTCTGCTGGGGCGTAATCACTCTCCTTAATGACAGACTG 1655

Db 1610 GAGTTTCACTGTACTGCGGAGATATGTGCTGCTCTTCTACGCGCCGCTCATGATGATCA 1668

QY 1656 TGTTCCTCTACGCTCTTCTCTGCGCTTACAACAACGTGGTGTGGAGACAACACTACGTATACC 1715

Db 1670 TATTCATATACTCGATGCGTGCACCATCGAACCGATCAATATATAGGCAACTGTAC---TTCC 1726

QY 1716 CGACTGCTGTTATGTTTCTGTGATACCTGATGTTATTTTGGGCATGACGTTTGCCAA 1775

Db 1727 CCGAAGCCGCGCAACATTTCCGGCTGGCGTGTGTCGCCATGGGTCGCCACAGTTTCCAC 1786

QY 1776 TTGGAATTTGATTTTCTTTGTACAAATACGTTACCGAACCTTACGCGACGATCAAGA 1835

Db 1787 TGTGGGGGCTTTGGTACATCTCGCGCCATCCACAAGGACGCTACTGTGAAAGCTCTGGAAG 1846

QY 1836 AAGGCTTCCATCCAAACCCATGAGGGGTGCCCGCTCGCCGACAGAGAGGTGAGAAATGGA 1895

Db 1847 CCTGCTGAACCCCAAGCGATCGCTGGGGTCTCTGCCAATCCGGAATTTGCGGCGCAATGGG 1906

QY 1896 TGCAGTTCAAGCTGGAACGGAACCTCTTACACAAGAAAGATAAA 1938

Db 1907 TGAATCTTACAGATCAGAGGCGAGCCCAAAAGCCGACGCAAG 1949

RESULT 6  
ABL07751  
ID ABL07751 standard; cDNA; 1992 BP

AC	ABL07751;
XX	
DT	26-MAR-2002 (first entry)
XX	

Drosophila melanogaster expressed polynucleotide SEQ ID NO 17735.

KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ss.

05 *Drosophila melanogaster*.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P

[illegible]XX  
DT 11-11-54 11:11 AM

XX 2001 252000 252

DR P-PSDB; ABB63648.  
XY

PT New isolated nucleic acid de  
PT from *Drosophila* and

PT interactions -  
VY

PS	Claim 1; SEQ ID NO 17735; 21
YY	

CC The invention relates to an  
CC capable of detecting 1000 or

CC useful in developmental biology

cc insecticides, therapeutics a

Claim 1; SEQ ID NO 17735; 21pp + Sequence Listing; English.













D6	754	ACCAAACTGGGAGTTGGTCAATTGGACTTTTATTCGGCTGGCTCTGTGTTTCAATAT	813
QY	820	CGCCCGGAGAGTCAAGAGTTCCGGCAAGCGGCTTACTTCTCGGCGCTCTCCATACGT	879
D6	814	TCGTCGGGAGTAGAAGTTCCGGGAAAGGCAATCTTCTCTG66CCCTTTTCCGACAT	873
QY	880	TGTCATATCATCTTATTCATCACACAAATCATCTCGCCGCGGTACTACAGGCAATCC	939
D6	874	CATCATGGGTCTTCTTGGTGGGAGCTGTATCTCTCCGGATTCATAATAGCGTATATA	933
QY	940	GTTCTTCGTACAGCGCTCAATGGCGAAATCTTGAGCTCGGTATGTACTACAGCAGT	999
D6	934	CTACTTATTAAGCCGCAATGGGGAAAGATTTGGACCCGAAAGTTTGGTACCGCGGTG	993
QY	1000	CACCCAGTGTCTTCTCTCGACAGCTGCACCGGACGAGATATATATGTTCTCTCTTA	1059
D6	994	AACGCACTGTTTCTACTCTCTGTCCGTTTCTTTTGGCAACATCATCTCATCTCTCT	1053
QY	1060	CAACGCTTTCAGACATATATCTCAAGAGATGCTTGATTTACAGACTTTGGACACTT	1119
D6	1054	CAACAATCTTGGGGACAAATGTACACAGAGATGACGAATTTGTAACGGGTCTGGACACT	1113
QY	1120	TACAAGTTTCTTGTCTGGGTGCACAGATCTTCGTATCTTGTATACCTCGCGTACGACT	1179
D6	1114	GACCTCTCTGTACTCTGATTCACAGATATTCGGTATCTCGGCAATCTGGCCACGAAT	1173
QY	1180	CAACTCAGAGTGGGAGATGTGGTGGCTGCGGGTACACAGATCTTTCATTTTCATA	1239
D6	1174	CGGAACCGATACATCGGCTCATGTTGTGAAGGGGGTCCGGANTGGCTTCATATCTATA	1233
QY	1240	CCGTGATGCCATTCGCCAA--ACATTTCAACCTCACTATTTCTCGGTCTCTTCTCT	1296
D6	1234	TCCCGATGCCATTCGCCAAATTCGCCAAATCTTCTCACTGCTGTCTCTCTCT	1293
QY	1297	GATGATGCGGTGCGGGGTATCGGCTCATCGTGGCTCTGCTATCGACTTTCACACATT	1356
D6	1294	CATCTCTCTCTGTGGGCAATAGATGGAATATAGCATATGACCTCGAGCTCGGACGCG	1353
QY	1357	GGCGATGAGCGCTTCCACGCTGTACCCACCGCTTACATGTCAAGCATGACCTTCTTG	1416
D6	1354	CATTGCGGATGCGTTTCCCACTTTGGGCACTGGCAAGTGCCTGCTCATCGCTGTAGT	1413
QY	1417	CGGTTTCCTGCTGGACTGTTTTCTGTCACCGGGTGGACAATATTTCTGACCTGT	1476
D6	1414	CAGTTTCTTTTATGATTAATGTATATTAACACCGGGCGGTAGATATGCTGACTCTGT	1473
QY	1477	AGATCACTACGCTGGACATCTCTGTGCTTTCTGCGCCATTTCTGAATCTCGACGCGT	1536
D6	1474	GGACTTTTTCGGCGCTCAATGATGTCTTGTGTACTGGGAATCCCGAGCTGTACACAT	1533
QY	1537	GTTCTGGAATTAATGATATGGAGAAATCTGTGCTAGACATTAAGTTCAATGTTGGGTAA	1596
D6	1534	TGGATGGATTAATGACACAGATCGCTGTGCAAGACATTAAGTTCATGCTGGGTCCGA	1593
QY	1597	GACTGTGCTTACTTGGCGTCTCTGCTGGGGGTATCACTCTCGTATATGAGACAGT	1656
D6	1594	AGTGGCGCTGTACTGGAGGCTGTGTGGAGATCATTAACCCGATCATGACTGTAT	1653
QY	1657	GTTCTTCTACGCTCTTCTCGCTCTAACAACTGTGTGGAGAACATCTAGTATACC	1716
D6	1654	TCTGATCTTA--TTTCTATGGACCTTCAACCTCTAATCAACAATATATGCTATACC	1710
QY	1717	GACTACCTGGTATATGTTCTGATCACTTATGATGTATTTTGGGCATGAGCTTGGCCAT	1776
D6	1711	AAATTGTCTTACTAATTTGGTGTATATAAGCGCTTCGTGATTTCTTACGTGCCAT	1770
QY	1777	TGGAATGGATTTCTTGTACAAATACCGTACCGGAACCTTCACGAGACATCAAGAA	1836
D6	1771	TTCGATGATCGTGGCCATTTGTGAGATCCAGGTACAGACTTGGGTGCAAAATCCGTGG	1830
QY	1837	AGCTTTCACATCCAAACCTCATGGGGTCCCG	1869
D6	1831	AGCCTTTACGCCGAAAGAAATTTGGGGACCCAG	1863

[illegible]

OY	960	GGGGAAACCTCCTTGAGGTGGGTGTATGGTCTCAGAGCATCCGACAGTGTCTTCTCTC	1019
Db	2142	GGCACAAGCTCTTGAAACCGCAGGTCTGTATGCGCGCGTACCCAGGTGTTCTTCTAC	2083
OY	1020	TGACAGTGTGCACCGGACCGCATCATATGTTCTCTCTTACAAAGGTTTTCAGACATATA	1079
Db	2082	TGGCCATCTGTTGGGAAACATCATATGTAAGGCTGTATACACCGCTTGGGCCAACA	2023
OY	1080	TCCTA-----	1087
Db	2022	TTTTCAGGTGAGTGGAGATATCCGGTTTGAGTTCTTACCTGTAGTTCTTCTCTCAGG	1963
OY	1088	GATACCTTGATTTGTATGACATTTTGAGACACCTTTCACAGTTTCTGTGCGGTGACATC	1147
Db	1962	GATCCCATATATCTGAGACACCGCTGGACACCTTTCACCTCCTACTGTCCGGTGTATATC	1903
OY	1148	TTGCGTATCCTTGGTAACTCGCGGTACGAACTCAGAGGTGGAGATGTGTGTGT	1207
Db	1902	TTTCGGATTTTCGGGCAACCTGGCTGTACGAGAAACAACACGACATCGTATGTTGTGTC	1843
OY	1208	GCTGGCGGTACCAATCTTGTCTTATTTATCATACCTGTATGCCATTCGCA--AACATTC	1264
Db	1842	AATGAGAGTCCGGGCGCTGGCTTTATATCTTACCGGACGCGCATTTGCAATTTTAACTGG	1783
OY	1265	CACCTCAGCTATTTCTCGGTCCTTCTTCTTCGTATGATGTGGGTGTGGGTATGGGTCA	1324
Db	1782	CTGGCGGAGATTTCTCGGTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1723
OY	1325	TCCGTGGCTCTGTATGACTTTCACACATTTGGCGATGAGCGGTTCCACGTTATCCC	1384
Db	1722	AATTTGGGCAAGGCGCTCTCGCATCTCCACCGCTGATCAAGGATCAGTTGCGACACCTGAAG	1663
OY	1385	ACCGTTCATCATGTCCAGGATACCTGTCTTTCGGGTTTCCGTTTGACACTGTATTACTGC	1444
Db	1662	AACCTGACCTGTGGTGTGGTATATGCCATTTTATGAGCTACTTTCGGGCTTCTTATCATTC	1603
OY	1445	ACACCGGGTGACATATATTTCTTGAGCTTTAGATCACTACGCGGTGACAACTTCTTGTG	1504
Db	1602	ACACCGGGTGCGCAGTTCCTGCTCAACCTGTGCACTACTTTGGCGTACAGTTGTGTGGCC	1543
OY	1505	CTTTTTCGCGCATTTCTGAACTCGCAGGCGTGTTCGTGATTTATGATTTGGAATCTG	1564
Db	1542	CTGTGTCTGTGCACTCTTCGACCTGGTGAACATTTCCCTGGAACTCAGGTGTGAACGCACTC	1483
OY	1565	TGCTTAGACATTTGAGTTTCATTTGGGTAAAGAACTGGTGTCTACTGGCGCTCTGCGTG	1624
Db	1482	TGCCGTGACCGGTGAGTTCAATGATTGGCATCAAGACTTCCGTACTATATCGCATCTGTGCG	1423
OY	1625	GGCGTAATCACTCTCTGCTATATATGACGACTGTGTCTTCTTCAAGCTCTTCTTCGCTTAAC	1684
Db	1422	GCGGTGTCACTCTCTGTCTATATGCTCACCATTCTCATCTTACACCTGGTGTCTGTACGAG	1363
OY	1685	AACCTGTGTTGGAGACAACTAGTATACCCGACTGCT	1723
Db	1362	CCCTCTCAAGTACAAGATTAACCTTACCAATCCGGGTGT	1324

RESULT 11	
ABL07750/c	
ID	ABL07750 standard; cDNA; 4385 BP.
XX	
AC	ABL07750;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 17732
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical; gene; ss.
XX	
OS	Drosophila melanogaster.
XX	
PN	W0200171042-A2.

XX 27-SEP-2001.  
PD  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
PF  
XX  
XX 23-MAR-2000; 2000US-191637P.  
PR  
PR 11-JUL-2000; 2000US-0614150.  
XX  
XX (PEKE ) PE CORP NY.  
PA  
PI Venter JC, Adams M, Li PWD, Myers EW;  
PI WPI; 2001-656860/75.  
DR P-PSDB; ABB63647.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
PT  
XX  
XX  
PS Claim 1; SEQ ID NO 17732; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB101840-ABL16175) and the encoded proteins  
CC sequences (AB101840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 4385 BP; 1367 A; 986 C; 884 G; 1148 T; 0 other;

Query Match	9.4%	Score 27.0	DB 23	Length 4385
Best Local Similarity	58.0%	Prod. No. 2.5e-49		
Matches	498	Conservative	0	Mismatches 358; Indels 3; Gaps 1;
QY	CTCAACTCTACTTTTGGAAACAGTTC	CCACAACAAAGGATGGAATTGAGAGGCTCG	755	
Db	2652	CGCAACTCTACTTGCACGACGATCG	GTGATGAAAGAGATTCCTCGAGGAGGCGATCG	2593
QY	756	GTGCCCCCATCTGGTACTTGGTGTGT	GTCTATTTCATTCATGCGCATGCGTCA	TGTTGGAG 815
Db	2392	GATATCCCAAGTGGAGTGGCTTAA	TGCGGGGCGATTTCGTGGCTAACCGTAACAC	TAA 2533
QY	816	TCGTCGCCGAGAGATCTCAAGATTC	CCGCGCAAGCGCCTACTCTCCGCGCTTCCAT	875
Db	2532	TCATCATTTCCGGGAGATGAAGATTC	GAGAAAGCGGCTATGTCTCGCCCTTTTCCGT	2473
QY	876	ACGTTGTCATGATCACTTATTTAT	CATCACACATCATCTCCCGGTC	TACTAGCGCA 935
Db	2472	ACGTTGTAATTTTATCTCGCTCG	GAGACATCACTATCAAGAGTCTCAACGAGAG	2413
QY	936	TCGCTTCCTTCGTACGCTCAATG	GGGCGGAACCTCTGACTCGCGTATGATCTAG	995
Db	2412	TTATGTACTTTCTGTGCTCTCAC	AGTGGGAACCTTTTGGACCCGAGGTGGTATCAATG	2355
QY	996	CAGTCAACGAAGTGTCTTCTCT	CTGCAAGTGTGACCGGACCGCATCTCATGTTCTCT	1055
Db	2352	CCGTCACCCAGAGTGTCTTCTT	CCCGGCGCTGTGTTTGGTGTGATCATCATATGATCTCT	2293
QY	1056	CTTACACAGGTTTCAGACATATAT	CTTACAGGATGCTTGGATTTTACGACTTTTGACA	1115
Db	2292	CGTCAACACCGTTGGTGTACAT	GTCTATCGGGAGCCCAATATATGTAACACCCCTGGACA	2233
QY	1116	CGTTTACAAAGTTCTTGTCTG	TGGGACAGATCTTCTGCGATCTCTTGGTAACCTCGGTACG	1175
Db	2232	CATTCACCTCACTGCTCTCG	GGGTGTGATTAATCTTTGGAACTCTGGAAACCTGGCTCAGC	2173
QY	1176	AATTCACATCAGAGGTGGAGAT	GTGTGCTGCTGCGCGTACCAAGTCTTGCTTTCATTT	1235

Db 2172 AATCGGGACCAAGACATTCAGTGTGTTAAGGAGGCTCCGTGATGGCTTTCATCT 2113  
Qy 1236 CATACCTGATGATGCAATGCAAA--ACATCCACCTGATCTGCTGCTGCTCT 1292  
Cc 1236 CATACCTGATGATGCAATGCAAA--ACATCCACCTGATCTGCTGCTGCTCT 1292  
Db 2112 CCAATCGGATGATGCAATGCAAAATGTTCTCCTCAAGTGTGCTGCTGCTCTCT 2053  
Qy 1293 TCCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1352  
Cc 1293 TCCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1352  
Db 2052 TCCCATGCTCTCATGTTGGGGGTGGTGAAGCATGAGGCAATGCTGATGATGTA 1993  
Qy 1353 CATGGCGATGAGCGGCTTCCACAGTACCCACCGCTACATGTAAGCATGCTGTT 1412  
Cc 1353 CATGGCGATGAGCGGCTTCCACAGTACCCACCGCTACATGTAAGCATGCTGTT 1412  
Db 1992 CCGTGTCAAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1933  
Qy 1413 CTGCGGTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1472  
Cc 1413 CTGCGGTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1472  
Db 1932 TGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1873  
Qy 1473 TGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1532  
Cc 1473 TGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1532  
Db 1872 TGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1813  
Qy 1533 GCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1551  
Cc 1533 GCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1551  
Db 1812 CCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1794

## RESULT 12

ABLI4420 standard; cDNA; 5270 bp.

ABLI4420:

26-MAR-2002 (first entry)

Drosophila melanogaster expressed polynucleotide SEQ ID NO 37742.

Drosophila; developmental biology; cell signalling; insecticide;

pharmaceutical; gene; ss.

Drosophila melanogaster.

W0200171042-A2.

27-SEP-2001.

23-MAR-2001: 2001MO-US09231.

23-MAR-2000: 2000US-191637P.

11-JUL-2000: 2000US-0614150.

(PEKE ) PE CORP NY.

Venter JC, Adams M, Li PMD, Myers EW;

WPI; 2001-656860/75.

P-PDB; ABB70317.

New isolated nucleic acid detection reagent for detecting 1000 or more

genes from Drosophila and for elucidating cell signalling and cell-cell

interactions -

Claim 1; SEQ ID NO 37742; 21bp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA sequences (ABLI1840-ABLI16175) and the encoded proteins

(ABBS7737-ABBS72072). The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO  
Cc at ftp.wipo.int/pub/published\_pot\_sequences.

Sequence 5270 bp; 1379 A; 1127 C; 1174 G; 1590 T; 0 other;

Query Match 8.7%; Score 247.4; DB 23; Length 5270;

Best Local Similarity 55.5%; Pred. No. 2.6e-44;

Matches 561; Conservative 0; Mismatches 391; Indels 59; Gaps 2;

Qy 749 GGTCTGGTCCCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 808  
Cc 749 GGTCTGGTCCCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 808  
Db 2932 GGGCTGGGATCTCCCAATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2991  
Qy 809 TTGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 868  
Cc 809 TTGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 868  
Db 2992 GGCACGATTTGTGAAGGCGCATTCGAGTTCGGAAGCACTTCTTCTGCGCCCTG 3051  
Qy 869 TTCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 928  
Cc 869 TTCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 928  
Db 3052 TTTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3111  
Qy 929 GACGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 988  
Cc 929 GACGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 988  
Db 3112 CAGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3171  
Qy 989 TATCAGCAGTCAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1048  
Cc 989 TATCAGCAGTCAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1048  
Db 3172 TACCGGCGCATCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3231  
Qy 1049 TTCTCTCTTACAAAGGCTTTCAGCATATAT----- 1078  
Cc 1049 TTCTCTCTTACAAAGGCTTTCAGCATATAT----- 1078  
Db 3232 TACCGGCGCATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3291  
Qy 1079 -----ATCAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1112  
Cc 1079 -----ATCAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1112  
Db 3292 TTTCAGATCATCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3351  
Qy 1113 ACACCTTTACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1172  
Cc 1113 ACACCTTTACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1172  
Db 3352 ACTCGCTTACATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3411  
Qy 1173 ACAGACTCAACTAGAGTGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1232  
Cc 1173 ACAGACTCAACTAGAGTGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1232  
Db 3412 TTGAGACCAACACCAAGATATTTCCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 3471  
Qy 1233 TTTCATACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1289  
Cc 1233 TTTCATACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1289  
Db 3472 TTTCATACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3531  
Qy 1290 TCTTCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1349  
Cc 1290 TCTTCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1349  
Db 3532 TCTTCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3591  
Qy 1350 ACACATGCGATGAGCGGCTTCCACAGTACCCACCGCTACATGTAAGCATGCT 1409  
Cc 1350 ACACATGCGATGAGCGGCTTCCACAGTACCCACCGCTACATGTAAGCATGCT 1409  
Db 3592 TGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3651  
Qy 1410 GTTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1469  
Cc 1410 GTTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1469  
Db 3652 CGATTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3711  
Qy 1470 AGCTTGTAGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1529  
Cc 1470 AGCTTGTAGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1529  
Db 3712 ATCTCGGACCTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3771  
Qy 1530 CAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1589  
Cc 1530 CAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1589  
Db 3772 TGGCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3831  
Qy 1590 GTTAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1649  
Cc 1590 GTTAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1649  
Db 3832 ATGTGAAGACCACTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3891





Best Local Similarity 54.1%; Pred. No. 8e-39;  
Matches 503; Conservative 0; Mismatches 415; Indels 12; Gaps 2;

```
OY 746 GGAGGTCTCGGTGCCCATCTGTACTGTGTGCTATTCATTCAGTGGCTCATG 805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1372 GAATATCTGTGTGATGATGAGTGGCCCTTGCCGTTGCTTTCTGGCCCTGGGTGATT 1431
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 806 GTGTTCGAGTCTGTGCCCGAGAGTCAAGAGTCCGCAAGCGGCTTACTTCTCCG 865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1432 GTATATGATGCGTGGCAAAAGAAATTAAGACATCAGAAAAGTGTACTTCACAGCC 1491
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 866 CTCTTCCCATACGTTGTGATGACATTTATTCATCACACATCATCTCCCGTGTCT 925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1492 ACCTTCCCTTATGCTGTGCTGTATCTCTCATTTGAGGGGTACACCTGCTGGAGCT 1551
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 926 ACTGAGGCAATCCGTTCTTGTGTCACGCTCATATGGGGAAATCTTGAGCTGGGTGA 985
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1552 GGAGCCGATCTGTGTACTTCTATCACTAAGTGGGAGAACTCACGATGCCAGGTG 1611
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 986 TGTACTCAGCAGTACGAGAGTGTCTTCTCTGACAGTGTGACCGGAGCCATCATC 1045
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1612 TGGAGAGATGAGGCACTCATATTTCTTCTCCTGTCTGCGGCTGGGGAGGGCTCATC 1671
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1046 ATGTTCCTCTTCAACGGTTTCAGACATATATCTACAGGATGCTTGGATTGTACG 1105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1672 ACTCTTCTTCTTACAACTAATTCATACAACTGATACAGGACACGTTAATTGTAAAC 1731
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1106 ACTTGGACACCTTTACAGTTTGTGTGGGTGGACAGATCTTGGATCTTGTGTAC 1165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1732 TGCACCAACAGTGCCTAGACTGTGCTGGGTTGTGTGATCTTCTGTCTATGGCTTC 1791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1166 CTCGCTACGAACTCACTCAGAGTGGAGATGTGTGCTGGCTGGCGGTACCAAGCTT 1225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1792 ATGGCCAAACGAGCGCAAGTCAACATGAGAAATGTG---GCTGACCAAGGGCCAGGCATT 1848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1226 GCTTTCATTTCATACCTGATGCAATGCTCAAAACATTCACCTCAGCTATTCGTGGTG 1285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1849 GCATTTGTGTTTACCCAGAGCCTTAACAGGCTGCTCTCTCCATCTGTGGGCCATTC 1908
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1286 CTGTCTTCCGATGATGTGGTGTGCTGGGATTCGGGCTCATCCGCTCTGCTATCAGACT 1345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1909 ATCTTTTCCGATGCTTCTACGCTTGGACTTGCACACCATGTTTGTACCATGAGAGCC 1968
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1346 TTCACACATGGGAGAGAGCGCTCCACGCTACCCACCGTCTACATGTACAGCATG 1405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1969 ATGTGACCTCCATCTCGGATGATTTCCCAAGATCTGGCGCACACAAAGCCTGTGTTC 2028
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1406 AC-----CTGTTCCTGGGTTTCTGCTGTGACTGTTTACTGCACACCGGGTGA 1456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2029 ACCCTGGGCTGTGCTCATCTGTCTCATTTATGGGCTTCCCAATGATCACACAGGTTGA 2088
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1457 CAATATATCTTGTAGCTGTAGATCACTACGCTGGAACATTCCTTGTGCTTTCTGCGCC 1516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2089 ATCTACATGTTTCAAGCTTGTGAGACACTATGCTGCTCTATGCTTGTTCATCATGCCC 2148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1517 ATTTCTGACCTGCGAGCGGTCTCTGATTTATGATTTGAGATGTGAGAACTGTGCTTAGACATT 1576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2149 ATATTTGAGCTGTGTCATCTCTATGTGTAGCGCTTGCAGAGTCTGTGAAGACATC 2208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1577 GAGTTCAITGTTGGCTAAAGACTGTGCTTACTGCGCTCTGCTGGGGCGTATCACT 1636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2209 GAGATGATGATTTGATTCAGCCCAACATTTTGTGAAGGTCTGCTGGGCTTGTGCACA 2268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1637 CCTGCTATTAAGAGACTGTGTCTTCTAC 1666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2269 CCGACCATTTTAACGTTTATCTTGTCTC 2298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: July 26, 2003, 00:00:40  
Job time : 637 secs

THIS PAGE BLANK (USPTO)





Email: [http://www.fruitfly.org/EST\\_estefruitfly.berkeley.edu](http://www.fruitfly.org/EST_estefruitfly.berkeley.edu)  
 hit genomic AEO03832: arm:2R [4615216,4876965]  
 estimated-cyto:45F4-46B13: 04/09/2001  
 Plate: AT.275 row: G column: 3  
 High quality sequence stop: 637.

## FEATURES

## source

1. 682  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 /clone\_1lb="At Drosophila melanogaster adult testes pote7"  
 /sex="male"  
 /dev\_stage="0-3 day old Ore-R males"  
 /lab\_host="Plates At.10-At.120: DH5-alpha. Plates  
 At.121-At.319: DH5-alpha Tona"  
 /note="Organ: ADULT testes. Vector: pOTB7; Site\_1: EcorI;  
 Site\_2: XhoI; The mRNA for the testis library was made  
 from testes and seminal vesicles hand dissected from 0-3  
 day old Ore-R males. RNA kindly provided by the lab of  
 Margaret Fuller. Sized fractionated cDNAs were directly  
 ligated into pOTB7. Plasmid cDNA library."

BASE COUNT 120 a 214 c 196 g 152 t  
 ORIGIN

Query Match 7.0%; Score 200.8; DB 12; Length 682;  
 Best Local Similarity 63.0%; Pred. No. 2,7e-25;  
 Matches 310; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 684 GTACGACGAGTGTCTACTTCTTTTGAAACAGTCTCCAAACAGGAGGAAATTG 743  
 DB 191 GAACCAACCTCGGGGAGTCTATTTCAGAACATCATCTCGCGAGAAAGTACATCG 250  
 QY 744 AAGAGAGTCTGGGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 803  
 DB 251 ATGATGACATCGATATCCACAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 310  
 QY 804 TGGTGTTCGAGTGTCTGCGCCGAGAGTCAAGAGTTCGCGCAAGCGGCTTCTCTCG 863  
 DB 311 TCATCCGCGGAGTATGTTTCAGAGGTGTGAAGAGTTCGCGCAAGCGGCTTCTCTCG 370  
 QY 864 CGCTCTCCCATACGTTGTATGATCATCTTATTCATCACCACATCATCTCTCGCGTG 923  
 DB 371 CCTCTCTCCCATACGTTGTATGATCATCTTCTCTGCGCGGAGCATCATCTCTCGCG 430  
 QY 924 CTATGACGACATCTGTTCTTCTGTCAGCGCTCAATGAGGAGAACTCTGAGCTGGTG 983  
 DB 431 CCTTCAGCGGCTGTCTGTTCTCTGCGCGGCAATGAGGAGAACTCTGAGCTGGTG 490  
 QY 984 TATGATACAGAGTCAAGAGTCTTCTCTGACAGTGTGACCGGACCGCATCA 1043  
 DB 491 TCTGTACAGCGCGCGTCAAGAGTCTTCTCTGACAGTGTGACCGGACCGCATCA 550  
 QY 1044 TCATGTTCTCTTCAAGAGTCTTCTGATCATATATCTACAGGAGTCTGATGTTTA 1103  
 DB 551 TCATGTTCTCTTCAAGAGTCTTCTGATCATATATCTACAGGAGTCTGATGTTTA 610  
 QY 1104 CGACTTTCGACACCTTCAAGTCTTCTGATGATGATGATGATGATGATGATGATG 1163  
 DB 611 CCAGCGTGGACACCTTCAAGTCTTCTGATGATGATGATGATGATGATGATGATG 670  
 QY 1164 ACCTGCGGTACG 1175  
 DB 671 ACCTGCGGTACG 682

RESULT 2  
 BM589455 704 bp mRNA linear EST 25-FEB-2002  
 LOCUS 1960048737041 A.gam.ad.cDNA.blood1 Anopheles gambiae cDNA clone  
 DEFINITION 1960048737041 5', mRNA sequence.  
 ACCESSION BM589455  
 VERSION BM589455.1 GI:18885316  
 KEYWORDS EST.

SOURCE  
 ORGANISM  
 African malaria mosquito.

ORGANISM  
 Anopheles gambiae  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;  
 Anopheles.

REFERENCE  
 1 (bases 1 to 704)

AUTHORS  
 Holt, R.A., Lin, J.-J., Murphy, S.D., Evans, C.A., Kraft, C.L., Charlalab  
 R., Collins, F.H., Venter, J.C. and Hoffman, S.L.

TITLE  
 JOURNAL  
 Celia Anopheles gambiae EST project  
 Unpublished (2002)  
 COMMENT  
 Contact: Holt R.A.  
 Celia Genomics  
 45 W. Gude Dr., Rockville, MD 20850, USA  
 Tel: 2404531351  
 Fax: 2404534580

FEATURES  
 source  
 Email: [holtra@celera.com](mailto:holtra@celera.com)  
 Plate: N001004A80 row: 0 column: 01  
 Seq primer: M13 Reverse.  
 location/Qualifiers  
 1. 704

/organism="Anopheles gambiae"  
 /strain="RSP-ST (Reduced susc. to Permethrin - std.  
 chromosome)"  
 /db\_xref="taxon:7165"  
 /clone="19600449705511"  
 /clone\_1lb="A.Gam.ad.cDNA.blood1"  
 /dev\_stage="Adult"  
 /lab\_host="DH10b"

/note="Vector: pSport1; Site\_1: SalI; Site\_2: NotI; Whole  
 adult mosquitoes (mixed sex) frozen on liquid nitrogen 24  
 hours after human blood feeding. cDNA inserts >500 bp  
 cloned directionally into pSport 1. Not 1 site is 3'.  
 Clones available through the Malaria Research and  
 Reference Reagent Resource Center ([www.malaria.mt4.org](http://www.malaria.mt4.org))"

BASE COUNT 131 a 220 c 185 g 168 t  
 ORIGIN

Query Match 5.7%; Score 163.6; DB 13; Length 704;  
 Best Local Similarity 53.9%; Pred. No. 7,7e-19;  
 Matches 359; Conservative 0; Mismatches 304; Indels 3; Gaps 1;

QY 223 CACGAGTAAAGAGGAGTACCGCCAGACGATATGATGATGATGATGATGATGATG 282  
 DB 29 CACGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 88  
 QY 283 GATGCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 342  
 DB 89 ACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 148  
 QY 343 GTACCAAGAAAGAGAGGAGTCTTCCGAGGAGTACATGATGATGATGATGATGATG 402  
 DB 149 GTTCAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 208  
 QY 403 CAAGCTGTGTACTTACTTAAAGTGTCTCTGAGCAATTCAGTTCAAGAAATCTGTAA 462  
 DB 209 ACGCCGATCTACTTACTTAAAGTGTCTCTGAGCAATTCAGTTCAAGAAATCTGTAA 268  
 QY 463 AGTTGTGCAATTCACCGGCAATGAAAGTACTGATAGTACGACAGCTCCGCGCGG 522  
 DB 269 GGTGTAGACACTTACCTGCAATGAGTGTGATGATGATGATGATGATGATGATGATG 328  
 QY 523 TTACATCTGTCTTACTACGTGTGATCTGTGCTCTGCTGCTGCTGCTGCTGCTGCTG 582  
 DB 329 CACGCTATACCTACTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 388  
 QY 583 CTTCAGAGCCACTCTTCATGAGGCTATTGTGACGCTGAGTGGAGAACTGCTACCTTC 642  
 DB 389 GTTCAATCCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 448  
 QY 643 AGATTCACACTT---GCTGATCAGTCAACATCAGTCAACATGATGATGATGATGATG 699  
 DB 449 TACGCGGCTGCTGATGATGCTTGGACGACGATGATGATGATGATGATGATGATGATG 508

QY 700 ACTTACTTTTGAGAACAGTTCCTCCACAAAGGATGGAATGAGAGGCTTCGGTGC 759  
D 509 TCTGTACTCCGCGAAAGGTAATGCGCGTACACTGACATGAGATGGTGGGATA 568  
QY 760 CCCCATGTGACTGCTGTGTGTCTATTCATGCGATGCTCATGCTGCGAGTCT 819  
D 569 TCCGAGTGGGCTGCTGCTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 628  
QY 820 CGCCGAGAGTCAAGAGTTCGCGCAAGCGGCTTCTGCGGCTTCGCAATCGT 879  
D 629 GATCAAGGCAATTAAGAGCTCCGCAAGGTTTCTGCTTCTGCTGCTGCTGCTGCT 688  
QY 880 TGTGAT 885  
D 689 CATCAT 694

## RESULT 3

CNS0101/c  
LOCUS  
DEFINITION

CNS0101 777 bp DNA linear GSS 14-JUN-2001  
Anopheles gambiae GSS SP6 end of clone 19M22 of NotreDamel library  
from strain PEST of Anopheles gambiae (African malaria mosquito),  
genomic survey sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AL149894.1 GI:7010373  
GSS.  
African malaria mosquito.  
Anopheles gambiae

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;  
Anopheles.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 777)  
Genoscope.  
Direct Submission  
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :  
BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr  
Web : www.genoscope.cns.fr)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

2 (bases 1 to 777)  
Roth,C.W., Brey,P.F., Ke,Z., Collins,F.H. and Weissenbach,J.  
Direct Submission  
Submitted (16-FEB-2000) BMDI, Institut Pasteur, 25, rue du Dr.  
Roux, Paris 75015, France  
This clone is from an A. gambiae BAC library provided by F.H.  
Collins and sequenced by Genoscope in collaboration with the  
Laboratory of Biochem. and Biol. Molec. of Insects, Institut  
Pasteur.

FEATURES  
source

1. 777  
Location/Qualifiers  
/organism="Anopheles gambiae"  
/strain="PEST"  
/db\_xref="taxon:7165"  
/clone="19M22"  
/clone\_1lb="Notredamel"  
/note="end : SP6"  
BASE COUNT 194 a 209 c 217 g 145 t 12 others  
ORIGIN

Query Match 5.3%; Score 151.6; DB 17; Length 777;  
Best Local Similarity 54.8%; Pred. No. 8.9e-17;  
Matches 311; Conservative 5; Mismatches 249; Indels 3; Gaps 1;

QY 740 ATTGAAGAGCTTCGGTGGCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 799  
D 743 ACTGCGGCTAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 684  
QY 800 CTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 859  
D 683 GTGCGCATGACMATCATGATGATGATGATGATGATGATGATGATGATGATGATG 624  
QY 860 CTCGCGCTCTCCCATACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 919  
D 623 CTGCGCCTCTCCCGTACGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 564

QY 920 GGTCTACTGACGGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 979  
D 563 GCGCGCGGTACGGCTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 504  
QY 980 GGTGTATGATACGACAGTACGACGACGACGACGACGACGACGACGACGACGACG 1039  
D 503 GACCTGAGAGGAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 444  
QY 1040 ATCATCATGCTTCTCTCTTACAAAGGTTTACAAATATATATATATATATATATAT 1099  
D 443 GTCATGTGCTTCTCTCTTACAAAGGTTTACAAATATATATATATATATATATATAT 384  
QY 1100 GTTACGACTTGTGACACCTTATACAAAGTTTCTGCTGCTGCTGCTGCTGCTGCTG 1159  
D 383 ATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 324  
QY 1160 GGTACCTGCGCTACGACATCACTACGACGACGACGACGACGACGACGACGACG 1219  
D 323 GCGCATCTGGGACATCTGAMAAACGAAACGACTACACCAAGGTGTAAAGCGGCGAGT 264  
QY 1220 AGCTGTGCTTCAATTCATACCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1276  
D 263 GCGCTAACGTTCATACCTACCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 204  
QY 1277 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1304  
D 203 TTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 176

## RESULT 4

LOCUS

BC026339 3118 bp mRNA linear HTC 08-APR-2002  
Homo sapiens, soluble carrier family 6 (neurotransmitter  
transporter, betaine/GABA), member 12, clone IMAGE:479451, mRNA.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BC026339.1 GI:20072875  
HTC.  
Homo sapiens.  
Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 3118)  
Strausberg,R.  
Direct Submission  
Submitted (02-APR-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK  
COMMENT

NIH-MGC Project URL: http://mgc.ncl.nih.gov  
Contact: MGC help desk  
Email: gqabs-remail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLT)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: http://www-shgc.stanford.edu  
Contact: (Dickson, Mark) mdick@stanford.edu  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/ILNLT at: http://image.llnl.gov  
Series: IRAK Plate: 32 Row: d Column: 18  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4507036  
this clone has the following problem: frame shifted.

FEATURES  
source

1. 3118  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="LocusID:6539"  
/db\_xref="taxon:9606"

```

/clone="IMAGE:4794351"
/issue-type="Brain, hypothalamus"
/clone_lib="NIH_MGC_96"
/ld_host="DH10B"
/notes="Vector: Bluescript"
BASE COUNT      636 a      885 c      832 g      765 t
ORIGIN

```

Query Match 5.28; Score 149.6; DB 11; Length 3118;

Best Local Similarity 46.38; Pred. No. 1.1e-16; Mismatches 814; Indels 37; Gaps 6;

```

Matches 735; Conservative 0;
221 GACACGACTTAAAGGCTGACCGCCAGACGATGATGATGTCACACAACTGAATTC 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 352 GACACGAAACGACGACGAGTGAAGATCGGGGCCAATGACACAAAGATGATTT 411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 281 TTGATGCTTCATCGCTACATCCGTCGTTGGGTAGACGTGAGGGGTTCCCTTATC 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 412 GTGCTGCTACGCGGGGAGATCATGGGCTGGCAATGCTGAGAGTTCCCTATCTC 471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 341 GCGACACGAAATGAGAGGCTTCCGTGTCAGTCAGTCAGTCAGTCAGTCAGTC 400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 472 TCGTACAAACGAGAGTGAAGCTTCTTATCCCTACTTCTTCTTCTTCTGTC 531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 401 GGCAGGCTGTACTACTAGATGTCTCGACATTCAGTTCAAGAACTCTGT 460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 532 GGCATCCCGGTTCTTCTGAGAGTGGCGTTGGGCCAATACACACCAAGGAGTCTC 591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 461 AAGATT---GTCATTTACCGGGCCATGAAAGTACTGATACGCTCAAGCTGCCGC 517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 592 ACAGCCGAGAGAGATCCGCCCTTCCAGGCAATGCTGCGCATCTGTGCTATC 651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 518 TGGGCTTACATCCGTCGTGCTAGTGTGCTGTCGTCTGTCATTTACTTACGT 577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 652 GAGTCATATTTGATGCTACTACATATATCTCTCTGCGGCTCTCTTACTCTTC 711
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 578 ATGAGCTTCAGGCGCACTCTCCATGCGGCTATTTGTGAGCGCTGAGGGAACTGGTA 637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 712 AGTCCTCTCACTTCTGAGCTGCCCTGAGACCGCAACCACTTTGGAACACAGACAT 771
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 638 CCGTCAGAT-----CCAACACTGTGTCATCATCTCAACAACTACCATATGTACC 688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 772 TGCACGAGCTTTTGAACCACTCAGAGCGCGCACTGACCCCATTTGAAATTTACC 831
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 689 AGCAGTCTCACTCTACTTTTGAAGACAGTTCTCCAAACAAAGCGATTTGAAGA 748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 832 TCACCTGTCAATGAAATCTGAGAGACGAGTCTGGCAATCACTGGGCACTAA---T 888
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 749 GGTCTGGTCCCGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 808
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 889 GACCTGGGCTCCCTGCGCTGGAGCTGGCCCTGCTGCTGCTGCTGCTGCTGCTG 948
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 809 TTGCGAGTCTGCGCGGAGAGTCAAGATTCGCGGCAAGCGGCTTCTCTCGGCTC 868
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 949 TATTTTGCATCTGGAAGGGGCTCAATCCACAGGCAAGGTGTTTATTTACAGCCAC 1008
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 869 TTCCCATAGCTGTGATGATCACTTTATTCATCAACCAATCTCTGCCGCTGCTACT 928
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1009 TTTCCTGACCTGATGCTTGTGATTTGCTGATCAGAGGTGTCACCCCTCCGGAGCTAC 1068
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 929 GAGGCTCTCTGCTTCTGTCAGGCTCAATGGGCAACTCCTTGAAGCTGCTGATG 988
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1069 CAGGGCACTCTACTACTTGAAGCCAGATTTTCCGCTCAAGGACCCCTCAGGTGTG 1128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 989 TATCAGCAGTCAAGCAAGTGTCTTCTCTGACAGTGTGACCGGACGAGATCATG 1048
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1129 ATGAGATCGGACCCGATCTTCTTCTTCCATCTGCGACGAGGCTCTCTGACAGCC 1188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1049 TTCTCTCTTCAACAGGTTTCAGACATATATCTACAGAGTCTTGGATTTGATCAGCT 1108
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1189 CTGGGAGCTACAAACAGATATCAACAACTGCTACAGAGACGATCGGCTCTGCTTC 1248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1109 TTGGACACCTTTACAACTTCTGCTGCTGGTGCACGATCTTGGGATCTTGGTACCTC 1168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 1249 CTGACAGTGTCCACACACTTTGTGGCTGGGCTTTGTGTCTC-----TCCATCTGGGCTTC 1304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1169 GCGTAGCACTCAACTCAGAGGTGGAGATGTGTGCTGGCTGACCATCTGTCT 1228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1305 ATGTCCCAAGAGAGAGGGGTGCCATTTCTGAAGTGTGCCAGTACGCTCTGGGCTGCC 1364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1229 TTCAATTTCATACCTCGATGTCATTTGCCAAACATTCACACCTGATCTCTGGTCTG 1288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1365 TTCTATGCTTCCCAAGGCTGTGACTATGATGCTCTTATCCAGCTGTGCTGCTGCTG 1424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1289 TTCTCTGATGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1425 TTCTTTATCATGCTCATATTTCTAGGCTGTGACAGCAGCATTTGCTGTGAGTGTCTG 1484
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1349 AACACATTTGGCATGACGCGTTC-----AGTGTACCCACCGCTTAC 1393
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1485 GTGACACCTTCACATAGACATGTCCCGAGCAGCTCCGAAAGACGGGCGGCGAGCTC 1544
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1394 ATGTACAGATGACCTGCTTCTGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1453
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1545 CTATCTCTACCATCGCCGCTGATGCTACTGATGAGGCTTTTCTGCTGACCGAGGCG 1604
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1454 GACAAATATATTTCTGAGCTGTAGATCACTACGCTGGAACATTCCT---TGTCTTTTC 1510
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1605 GGGATGTACATTTCTCAGCTGTGTTGACTACTATGCTTCCAGTGGCATATGCTGCTGTC 1664
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1511 TGGGCACTTTCTGACATCGACGAGGCTTCTGATTTATGATTTGGAATCTGTGCTTA 1570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1665 CTGATGATTTTGAAGGTGTGATTAAGTGTGATGAGTGTGAGGCGGACGCTTTCTATGAC 1724
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1571 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1630
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1725 AACATTTAGGACATGATTTGGCTACCGGCAATGCGCCCTGGTGAAGATCTCTGCTCTTC 1784
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1631 ATCAGCTCTCTATATGACAGCTGTGCTTCTTCTGCTCTTCTGCTTCAACACCTG 1690
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1785 CTGACCCCTGAGCTTCTGCTGCGCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1844
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1691 GTGTGGAGACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1750
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1845 AAGTACAAACAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1751 TTTTGGGCACTGACGTTTGTGCCAAT 1776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1905 CTGCTCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 5  
Bg780471 798 bp mRNA linear EST 20-MAY-2001

LOCUS  
SEAM0000428 Sea urchin primary mesenchyme cell cDNA library

DEFINITION  
Strongylocentrotus purpuratus cDNA clone

ACCESSION  
Bg780471.1 GI:14151484

VERSION  
Bg780471.1

KEYWORDS  
EST.  
Strongylocentrotus purpuratus.

SOURCE  
Strongylocentrotus purpuratus

ORGANISM  
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
Echinoidea; Euechinoidea; Echinacea; Echinoida;  
Strongylocentrotidae; Strongylocentrotus.

REFERENCE  
Zhu,X., Mahlras,G., Illies,M.R., Cameron,R.A., Davidson,E.H. and  
Ettensohn,C.A.

A large scale analysis of mRNAs expressed by primary mesenchyme  
cells of the sea urchin embryo

Development 128 (13), 2615-2627 (2001)

CONTACT: Ettensohn CA  
Dept. Biol. Sci.  
Carnegie Mellon University  
4400 Fifth Avenue, Pittsburgh, PA 15213, USA

Tel: +1 412 268 5849  
Email: etlensohn@andrew.cmu.edu.  
Location/Qualifiers  
1. .798

FEATURES  
source  
/organism="Strongylocentrotus purpuratus"  
/db.xref="taxon:7668"  
/clone="91222951\_N20\_072\_PC\_0020\_B2\_MR\_G10"  
/clone\_lib="Sea urchin primary mesenchyme cell cDNA library"  
/tissue\_type="embryo"  
/cell\_type="primary mesenchyme cells"  
/lab\_host="E.coli"  
/note="Vector: pSPORT1, Site\_1: NotI; Site\_2: SalI; oligo dt priming from poly A+ RNA, directionally cloned"  
BASE COUNT 136 a 241 c 184 g 236 t 1 others  
ORIGIN

Query Match 5.1%; Score 145.2; DB 12; Length 798;  
Best Local Similarity 52.4%; Pred. No. 1.1e-15;  
Matches 374; Conservative 0; Mismatches 328; Indels 12; Gaps 2;

QY 848 GCGGCTACTCTCTCGGCTCTCCATACGTTGCATGATCACTTATTCATCCACACA 907  
DB 42 GTGGTACTTCAACAGCTACCTTCCCTTACCTTGTCCGTGATCCTTGTGCTGGA 101  
QY 908 ATCAATCTGCGGCTGCTACTGACGCGATCTGTTCTTGTACGCTCAATGGCGGAA 967  
DB 102 GTTACTTACCGGGGTCTTATGATGATGATTTGTTCTTCACTGCTGCTTACTACA 161  
QY 968 CTCTTGTAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1027  
DB 162 CTCAGCAACCTCTAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 221  
QY 1028 TGACCGGACGATCATCATCATCTCTTACACAGGTTTCAACATATATATACAGG 1087  
DB 222 GCTTGGGAGGATCTATCAACCTCTCATCTTACCAAGATTCACCAACATGCTACTT 281  
QY 1088 GATCTTGTAGTGTAGGATCTTGTGACACCTTCAAGTTTCTGCTGCTGACGATC 1147  
DB 282 GATCTTGTAGTGTAGGATCTTGTGACACCTTCAAGTTTCTGCTGCTGCTGATC 341  
QY 1148 TTGCTTGTAGTGTAGGATCTTGTGACACCTTCAAGTTTCTGCTGCTGCTGATC 1207  
DB 342 TTCTCTATCTGCTGCTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 401  
QY 1208 GCTTGGGATGACGATCTTGTGCTTCAATTCATGCTGATGCTGCTGCTGCTGCTG 1267  
DB 402 TCAGGATTT---GGTCTTGTGCTTCAATTCATGCTGATGCTGCTGCTGCTGCTG 458  
QY 1268 CTTAGCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1327  
DB 459 TCACCGGCTGCGGATCTCTCTTCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTG 518  
QY 1328 GTGCTCTGCTATGCTGCTTCAACATTTGCGATGCGATGCGATGCGATGCGATGCG 1384  
DB 519 TTACCATCATGAGAGAGGCTGCTGACAGCCCTGCTGAGAGTTCCCGACACCTCCG 578  
QY 1385 -----ACGCTTCAATGCTGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1438  
DB 579 AAGAAGAAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 638  
QY 1439 TACTGCAACCGGCTGCAATATATTTCTGAGCTTGTAGATCATACGCTGGAACATTC 1498  
DB 639 TGCATCAACAGAGGCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 698  
QY 1499 CTTTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1552  
DB 699 TCCCTCATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 752

RESULT 6  
BF486192 670 bp mRNA linear EST 19-APR-2001  
LOCUS BF486192

DEFINITION AT20411.5:prime AT Drosophila melanogaster adult testes POTB7  
Drosophila melanogaster cDNA clone AT20411.5 similar to CG3252;  
F800003252 located on: X 4F8-4F8; 04/09/2001, mRNA sequence.  
BF486192  
VERSION BF486192.2 GI:13698079  
KEYWORDS EST.  
SOURCE fruit fly.  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 670)  
Stapleton, M., Brokstein, P., Hong, L., Agapay, A., Baxter, E., Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Miranda, A., Mista, S., Mungall, C.J., Nuno, J., Pacלב, J., Paragas, V., Park, S., Phoumenyong, S., Wan, K., Yu, C., Lewis, S.E., Cealiker, S., and Rubin, G.M.  
BDGP/HMI AT Drosophila EST Project  
Unpublished (2000)  
On Dec 6, 2000 this sequence version replaced gi:11569493.  
Contact: Stapleton, M.

TITLE BDGP  
JOURNAL Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: <http://www.fruitfly.org/EST>, [est@fruitfly.berkeley.edu](mailto:est@fruitfly.berkeley.edu)  
hit genomic AB003434; arm: X [5133852,5435637]  
estimated-cyto:4F2-5A9; 04/09/2001  
Plate: AT 204 row: A column: 11  
High quality sequence stop: 565.  
location/Qualifiers

FEATURES  
source  
1. .670  
/organism="Drosophila melanogaster"  
/db.xref="taxon:7227"  
/clone="AT20411"  
/clone\_lib="AT Drosophila melanogaster adult testes POTB7"  
/sex="male"  
/dev\_stage="0-3 day old Ore-R males"  
/lab\_host="Plates AT 10-AT 120; DHS-alpha. Plates AT 121-AT 319; DHS-alpha TomA"  
/note="Organ: ADULT testes; Vector: POTB7; Site\_1: EcoRI; Site\_2: XhoI; The mRNA for the testis library was made from testes and seminal vesicles hand dissected from 0-3 day old Ore-R males. RNA kindly provided by the lab of Margaret Fuller. Sized fractionated cDNAs were directly ligated into POTB7. Plasmid cDNA library."  
BASE COUNT 165 a 155 c 190 g 160 t  
ORIGIN

Query Match 5.0%; Score 141.8; DB 12; Length 670;  
Best Local Similarity 58.4%; Pred. No. 4.8e-15;  
Matches 248; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 218 GATGACGAGCTTAGAGGCTGGAACCGCCAGACGTATGATGCTCCACACATTTGAA 277  
DB 141 GAAAAACAGATGCGCAGAACCGCAGCGAGCGACCAATTTGGGCAATGCGCTGAG 200  
QY 278 TTCTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 337  
DB 201 TTCTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 260  
QY 338 ATCGCTTACAGAAATGAGAGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 397  
DB 261 ACGGCTTATGAGAAATGAGAGAGGCGCTTCTTATACCTTACATCATGATGCTCTTTC 320  
QY 398 GTGCGCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 457  
DB 321 ATCGCAAAACCATGATCTATTTGAGATGATGATGATGATGATGATGATGATGATGAT 380  
QY 458 GTTAAAGTTTGTGCTAATTTACCGGCGCATGAAGTACTGATAGCTGCAAGCTGCGCG 517  
DB 381 GTGAGATTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 440

[illegible][illegible]







LOCUS	BM653918	707 bp	mRNA	linear	EST-26-FEB-2002
DEFINITION	17000687379138 A.Gam.ad.cDNA1 Anopheles gambiae cDNA clone				
ACCESSION	196004496619090 5', mRNA sequence.				
VERSION	BM653918				
KEYWORDS	BM653918.1 GI:18953429				
SOURCE	EST.				
ORGANISM	African malaria mosquito.				
	Anopheles gambiae				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;				
REFERENCE	1 (bases 1 to 707)				
AUTHORS	Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab				
TITLE	R., Collins,F.H., Venter,J.C. and Hoffman,S.L.				
JOURNAL	Celera Anopheles gambiae EST project				
COMMENT	Unpublished (2002)				
	Contact: Holt R.A.				
	Celera Genomics				
	45 W. Gude Dr., Rockville, MD 20850, USA				
	Tel.: 2404533151				
	Fax: 2404534580				
	Email: HoltRA@celera.com				
	Plate: ND01004NBT row: N column: 16				
	Seq primer: M13 Reverse.				
FEATURES	Location/Qualifiers				
source	1..707				
	/organism="Anopheles gambiae"				
	/strain="RSP-ST (Reduced susc. to Permethrin - std.				
	chromosome)"				
	/db_xref="taxon:7165"				
	/clone="19600449619090"				
	/clone_lib="A.Gam.ad.cDNA1"				
	/dev_stage="Adult"				
	/lab_host="DHI0b"				
	/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI. Whole				
	adult mosquitoes (mixed sex) frozen on liquid nitrogen.				
	cDNA inserts >500 bp cloned directionally into pSPORT 1.				
	Not 1 site is 3'. Clones available through the Malaria				
	Research and Reference Reagent Resource Center				
	(www.malaria.mr4.org)."				
BASE COUNT	141 a 187 c 189 g 190 t				
ORIGIN					
Query Match	4.4%; Score 126.4; DB 13; Length 707;				
Best Local Similarity	59.4%; Pred. No. 2.2e-12;				
Matches 214; Conservative	0; Mismatches 146; Indels 0; Gaps 0;				
OY	739	AATTGAAGAGAGTCTCGGTGCCCCCATCTGGTACTTGTTGTTGTCATTCATCGGATG	798		
DB	335	ATTGGCAGAGGAGACTTGATGGCCAGATGGCAAGTGGTCTCTGTCTACTCGTCTCTG	394		
OY	799	GCTCATGGTGTTCGGAGTCTGTGCCCGGAGATCAAGAATTCCGGCAAAGCGGCCTACTT	858		
DB	395	GTCCATTCGTGTTGGATCTCTATCAAAAGGATTTGCGACACCGGGAAGCGGCTACTT	454		
OY	859	CCTGGCGCTCTCCCATACGTTCGATCATATCTATTATTCATCACCAACAATCAATCCCTC	918		
DB	455	TCTTGCCACTCTCCGTAATGTAATCATTTTCACTCTGCTGCCCACTCCCTCAGTCTGA	514		
OY	919	CGGTGCTACTGACGGCATCTCTGTTCTTCGTACACGGCTCAATGAGGGAACCTCTTAGCT	978		
DB	515	AGGTTCCGCTCAGAGGATATCAAGTTCTTCTTAACCCCCAAAATGGGAACCTATTCCTGAC	574		
OY	979	CGGTGATNGTACTCAGCAGTCAGCAAGTGTCTCTCTGTACAGTGTGCACCAGCC	1038		
DB	575	CAAGTATGATGAGGAAGCAGTCACACACATGTCTTCTTCCCTGTCAATCTTTTGGGAG	634		
OY	1039	GATCATCATCTCTCTCTTACACAGGTTTGAGACATATATCTACAGGATGCTTGAT	1098		
DB	635	AATCATTTGCGTACTCTCGTTCCAATAATTTTCCCATATATGTCTATGTGATGCAATGAT	694		

LOCUS	BM581199	589 bp	mRNA	linear	EST 22-FEB-2002
DEFINITION	I7000687244048 A.Gam.ad.cdNA.blood1 Anopheles gambiae cDNA clone				
ACCESSION	1960044963580 5', mRNA sequence.				
VERSION	BM581199				
KEYWORDS	BM581199.1 GI:18869666				
SOURCE	EST.				
ORGANISM	African malaria mosquito.				
REFERENCE	Anopheles gambiae				
AUTHORS	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.				
TITLE	1 (bases 1 to 589)				
JOURNAL	Holt,R.A., Iln.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab				
COMMENT	R., Collins,E.H., Venter,J.C. and Hoffman,S.L. Celera Anopheles gambiae EST project Unpublished (2002) Contact: Holt R.A. Celera Genomics 45 W. Gude Dr., Rockville, MD 20850, USA Tel.: 2404533151 Fax: 2404534580 Email: HoltR@celera.com Plate: MD01004B2V row: M column: 06 Seq primer: M13 Reverse. Location/Qualifiers 1..589 /organism="Anopheles gambiae" /strain="RSP-ST (Reduced susc. to Permethrin - std. chromosome)" /db_xref="taxon:7165" /clone="19600449683580" /clone_11b="A.Gam.ad.cdNA.blood1" /dev_stage="Adult" /lab_host="DH10b" /note="Vector: pSport1; Site_1: SalI; Site_2: NotI; whole adult mosquitoes (mixed sex) frozen on liquid nitrogen 24 hours after human blood feeding. cDNA inserts >500 bp cloned directionally into pSport 1. Not 1 site is 3'. Clones available through the Malaria Research and Reference Reagent Resource Center (www.malaria.mr4.org)"				
FEATURES	BASE COUNT 106 a 173 c 150 g 160 t				
ORIGIN	Query Match 4.4%; Score 124.8; DB 13; Length 589; Best Local Similarity 52.3%; Pred. No. 4.5e-12; Matches 301; Conservative 0; Mismatches 272; Indels 3; Gaps 1;				
	<p>1030 CACCGGACCGATCATCATGTTCTCTCTTAAACGGTTTCAGACATATATCTACAGGA 1089</p> <p>1090 TCGTTGAGTTCTTACAGCTTTGGACACCTTTACAAGTTCTTGTCTGGGTCACAGATCT 1149</p> <p>1150 CGGATTCCTTGGTAACCTCCGCTACGACACTCACTCAGAGTGGAGATGTGGTCGGTC 1209</p> <p>1210 TGGCGGTACCAAGTCTTGGCTTCATTTCAATTCACCTCGATGCAATTCGCA 1266</p> <p>1267 AACTGAGTATTTCTCGGTGCTGTTCTTCCTGATGATGCTGGTCTGGATGGCTATC 1326</p> <p>1327 CGTGGCTCTGATGACACTTTCACACACTTGGCAGTGGAGCGCTTCACAGTACCCAC 1386</p> <p>1387 CGTCACATGTCAGCGATGACCTGTTCTTGGCGTTTCCCTGCTGGATCTGTTACTGCAC 1446</p>				

Db 363 ATGGAAGATGATGATCGGGTCCGAGGAGGCTTGCATTCGATGGTTCATCAGAC 422  
QY 1447 ACCGGGAGCAATATATCTTACGCTTGTAGATCACTAGCGAGCAATTCCTGTGCT 1506  
Db 423 ACCGGGAGGTTTATCTGATCGAGCTGTGATTAATCAAGCGCCGCTGGTACCAT 482  
QY 1507 TTTTCGCCCATTTCTGACACTCGAGGCTGTCTGTGATTTATGATTTGAGAAATCTGTG 1566  
Db 483 AAGCGTGGCGCTGTGACGCTGCACCTTTGCTGATCTATGCGCTGATCGGCTGTG 542  
QY 1567 CCTAGACATTCATCTCTGTTGGTAAAGAACTGG 1602  
Db 543 CAAGACATTCATGATTCCTCGCATCAAGACAGG 578

RESULT 14  
BM963831  
LOCUS  
DEFINITION  
BM963831 777 bp mRNA linear EST 18-MAR-2002  
UI-M-EQ0-bwn-j-09-0-UI.r1 NIH\_BMAP\_E00 Mus musculus cDNA clone  
IMAGE:5698664 5', mRNA sequence.  
BM963831  
BM963831.1 GI:19547251  
EST  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 777)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Dr. James Liu, University of Iowa  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL at:  
http://image.llnl.gov  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)

FEATURES  
source  
Location/Qualifiers  
1..777  
/organism="Mus musculus"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5698664"  
/clone\_1lb="NIH\_BMAP\_E00"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 15.5 dpc"  
/lab\_host="DH10B (TI phage resistant)"  
/note="Organ: brain; Vector: pYX-Asc; Site: 1; Ecor I;  
Site: 2; Not I; The library was constructed according to  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured mRNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with an  
oligo-dT primer containing a Not I site. Double stranded  
cDNA was size selected according to mRNA size fraction  
ligated with Ecor I adaptor, digested with Not I, and then  
cloned directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
, is CAGCCAGCAGC. This library was created for the  
University of Iowa Mouse Brain Molecular Anatomy Project  
(BMAP). 'Gene Discovery in the Developing Mouse Nervous  
System', supported by National Institutes of Mental Health  
(NIMH), Hemlin Clin, Ph.D., program coordinator."

BASE COUNT  
134 a 233 c 193 g 215 t 2 others

Query Match 4.2%; Score 121.4; DB 14; Length 777;

Best Local Similarity 50.8%; Pred. No. 1.6e-11;  
Matches 370; Conservative 0; Mismatches 348; Indels 11; Gaps 3;  
QY 766 CTGCTACTGCTGTGTTGTTCTATTCATGCATGAGCTCATGTTGGAGTCTGCCCCG 825  
Db 24 CTGCGCTTGGCCATCATCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 83  
QY 826 AGAGCTGAAGATTCGGGCAAGGGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 885  
Db 84 GGGTGTGTTGGTGGAGCTGGAAAGGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 143  
QY 886 GATCATTATTCATCATCACAATCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 945  
Db 144 TATCATCTGTTCTTCTGCTGAGTACGCTTCCGGGGCCAAAGAGGGATTCCTCTCTA 203  
QY 946 CGTCAGCCCTCATAGGGCGAAAGTCTCTGAGCTGCTGCTGCTGCTGCTGCTGCTG 1005  
Db 204 CATCACAACCCCACTCCGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 263  
QY 1006 AGTGTCTCTCTCTGACAGTGTGCAACCGAGCATCATCATGCTCTCTTACAAAGG 1065  
Db 264 GATCTTCTCTCTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 323  
QY 1066 TTTTCAGACATTAATATCTACAGGATGCTGATTTGATACGACTTGGACACTTACAG 1125  
Db 324 TTTCCACACAAATGTGACAGGATCATCATCTGTTGCTGATCAACTCTGACACAG 383  
QY 1126 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1185  
Db 384 CATGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 443  
QY 1186 AGAGTGGAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1245  
Db 444 GTCCATATGCTGATGTGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 500  
QY 1246 TGGCATTTGCCAATAATTCACACCTGATTTCTGCTGCTGCTGCTGCTGCTGCTG 1305  
Db 501 GGGTGTGACACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 560  
QY 1306 GGTGCTGGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1365  
Db 561 GATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 620  
QY 1366 CAGCTTCCACAGTATACCCACCTGCTACATG-----TCAGCATGACCTGTTCTGCGG 1419  
Db 621 CGAGTACCCAGACTTCCGCAATGCGCGAAGCTCTTATTCCTGCTGCTGCTGCTGCT 680  
QY 1420 T--TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1477  
Db 681 TCCTACTCTATTTGGCTGTCTACATCACCACCAAGGTGCAATTAATGCTTCANACGTGTT 740  
QY 1478 GATCACTAC 1486  
Db 741 GATTATTAC 749

RESULT 15  
CNS01RBB 584 bp DNA linear GSS 14-JUN-2001  
LOCUS  
DEFINITION  
CNS01RBB Anopheles gambiae GSS T7 end of clone 32110 of NotreDame library  
from strain PEST of Anopheles gambiae (African malaria mosquito),  
genomic survey sequence.  
AL157191  
AL157191.1 GI:7018110  
GSS.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Anopheles gambiae  
African malaria mosquito.  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;  
Anophelinae.  
1 (bases 1 to 584)  
Genoscope.  
Direct Submission

REFERENCE  
AUTHORS  
TITLE



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2003, 23:43:13 ; Search time 131 Seconds

(without alignments)  
6695.386 Million cell updates/sec

Title: US-09-991-458-1

Perfect score: 2860

Sequence: 1 ggcacgaggtactactgtgtg.....ataaaaaaaaaaaaaaaaaa 2860

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

1: /cgn2\_6/prodata/1/lna/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/lna/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/lna/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/lna/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/lna/PCBUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/lna/backfiles1.seq:\*

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	222	7.8	2403	2	US-08-700-013B-26
2	222	7.8	2817	1	US-08-655-836-4
3	222	7.8	2817	2	US-09-020-753-4
4	209.6	7.3	2728	1	US-07-879-617A-7
5	209.6	7.3	2728	1	US-08-753-985-7
6	209.4	7.3	3404	1	US-07-762-132A-1
7	204.8	7.2	3919	1	US-08-301-722A-1
8	196.4	6.9	2394	4	US-09-191-468-121
9	196.2	6.9	2278	1	US-07-959-943-6
10	196.2	6.9	2415	1	US-07-959-943-8
11	194.8	6.8	2394	4	US-09-191-468-119
12	193.2	6.8	2394	4	US-09-191-468-123
13	193.2	6.8	2397	2	US-08-700-013B-18
14	193.2	6.8	2397	2	US-08-700-013B-20
15	191.6	6.7	2863	4	US-09-182-718A-1
16	191.6	6.7	2863	4	US-09-795-232-1
17	189.4	6.6	2756	1	US-07-782-298-1
18	187.2	6.5	2508	1	US-07-959-943-10
19	187.2	6.5	2064	3	US-08-834-467-1
20	161.4	5.6	2064	4	US-09-336-177-1
21	161.4	5.6	2136	1	US-08-655-836-1
22	161.4	5.6	2136	2	US-09-020-753-1
23	161.4	5.6	2169	3	US-08-834-467-3
24	161.4	5.6	2169	4	US-09-396-177-3
25	161.4	5.6	2202	1	US-08-655-836-2
26	161.4	5.6	2202	2	US-09-020-753-2
27	161.4	5.6	2364	1	US-08-655-836-3

28	161.4	5.6	2364	2	US-09-020-753-3	Sequence 3, Appl1
29	160	5.6	1991	4	US-08-295-814E-9	Sequence 9, Appl1
30	160	5.6	1991	4	US-09-343-361-9	Sequence 9, Appl1
31	160	5.6	1991	5	PCT-US93-01959-9	Sequence 9, Appl1
32	156.2	5.5	1938	1	US-08-295-814E-3	Sequence 3, Appl1
33	156.2	5.5	1938	4	US-09-343-361-3	Sequence 3, Appl1
34	156.2	5.5	1938	5	PCT-US93-01959-3	Sequence 3, Appl1
35	154.8	5.4	2217	1	US-08-543-881-1	Sequence 1, Appl1
36	154.8	5.4	2217	1	US-08-291-299-1	Sequence 1, Appl1
37	154.8	5.4	2217	5	PCT-US94-00119-1	Sequence 1, Appl1
38	154.8	5.4	2217	5	PCT-US95-10579-1	Sequence 1, Appl1
39	151.6	5.3	2121	1	US-08-240-783B-1	Sequence 1, Appl1
40	151.6	5.3	2121	3	US-09-084-813-1	Sequence 1, Appl1
41	151.6	5.3	2121	5	PCT-US97-09662-1	Sequence 1, Appl1
42	148.4	5.2	755	4	US-09-191-468-71	Sequence 71, Appl1
43	147.6	5.2	2028	1	US-08-295-814E-1	Sequence 1, Appl1
44	147.6	5.2	2028	4	US-09-343-361-1	Sequence 1, Appl1
45	147.6	5.2	2028	5	PCT-US93-01959-1	Sequence 1, Appl1

#### ALIGNMENTS

RESULT 1

US-08-700-013B-26

Sequence 26, Application US/08700013B

Patent No. 5919553

GENERAL INFORMATION:

APPLICANT: Albert, Vivian R.

APPLICANT: Kowalski, Leslie R.Z.

APPLICANT: Bordone, Laurence A.

APPLICANT: McKelvey, Jeffrey F.

TITLE OF INVENTION: Human Glycine Transporter

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESSES:

ADDRESS: Dechert Price & Rhoads

STREET: 997 Lenox Drive, Building 3, Suite 210

CITY: Lawrenceville

STATE: NJ

COUNTRY: USA

ZIP: 08543

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTED for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/700,013B

FILING DATE:

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Bloom, Allen

REGISTRATION NUMBER: 29,135

REFERENCE/DOCKET NUMBER: 317743-108

TELECOMMUNICATION INFORMATION:

TELEPHONE: 609-520-3214

TELEFAX: 609-520-3259

TELEX:

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 2403 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-700-013B-26

Query Match: 7.8%; Score 222; DB 2; Length 2403;

Best Local Similarity 54.1%; Pred. No. 1,1e-44;

Matches 503; Conservative 0; Mismatches 415; Indels 12; Gaps 2;

```

QY 746 GGAGTCTCGGTGCCCCCATCTGTACTGTGTGTATTCATCGATGGCTCATG 805
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1165 GAATATCTGTGTAGATCAGATGGCCCTTCCGCTTTTCCGGCTGGGTAT 1224
QY 806 GTGTGGAGTCTGCGCCGAGATCAGAGTTCGGGCAAGCCGCTACTCTCGCG 865
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1225 GTATATGCAATCGCTGGGAAAGATTAAGACATCAGGAAAGGTGTACTTACAGCC 1284
QY 866 CTCTCCCATAGCTGTGATGATCATCTTTATTCATCCACCAATCCGCGGCTGCT 925
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1285 ACCCTCCCTTATGTCTGTCTGTATCTCTCATTCGAGGGGTACCTGCTGGAGCT 1344
QY 926 ACTGAGGACCTCTGTCTGTCTGTACGCTCAATGGGCGAAACTCTTGTAGCTGCTGTA 985
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1345 GGAGCCGGTATCTGTACTCATCATCAGCTTAAGGGGAAACTCAGCGATGGCAAGCTG 1404
QY 986 TGTACTCAGCATCTCAGCAGAGTGTCTTCTGTGACAGTGTGACCGGACCGATCATC 1045
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1405 TGGAGGATGCAACCTCAGATTTCTTCTCCCTGTCTGCGGCTGGGAGGGCTCATC 1464
QY 1046 ATGTCTCTCTTACAAAGCTTTCAGACATATATATACAGGGATGCTTGATTTAG 1105
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1465 ACTCTTTCTTCTTACAAATTCATACACTGCTACAGGACGTTAATTTGTAAC 1324
QY 1106 ACTTTGACACCTTTACAACTTTCTGTCTGTGTCAGATCTCTGGTATCTTGTATAC 1165
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1525 TGCACCAACAGTCCCATACATCTTCTGTGGTTGTCTCTCTCATTTGCTGATGGCTTC 1584
QY 1166 CTGGCGTACGAATCAATCAGAGTGGGAGATGTGTGTGGTGTGGCTGACAGTCTT 1225
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1585 ATGGCCACAGAGGCAAGTCAACATTTGAGAAATGTG--GCTGACCAAGGGCGAGGATTT 1641
QY 1226 GCTTTCATTTCAATACCTGTATGCTGATTCACCAAAATTCACCACTAGCTATCTCGGTG 1285
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1642 GCATTTGTGTATACCAAGAGGCTTACCAAGGCTGCTCTCTCATTTCTGGGCTATC 1701
QY 1286 CTGTCTCTCTGATGATGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1345
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1702 ATCTTTTCTGTATGCTTCTCAGCTGTGACTTGACACCAATGTTGTACTACATTCAGAAC 1761
QY 1346 TTCAACATATGGGAGTGGGAGGCTGCTCCACGCTGTACCCGCTGTACATGTACAGCATG 1405
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1762 ATGTGACCTTCATCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1821
QY 1406 AC-----CTGTCTTCTGCGTGTCTGCTGTGACTTGTATCTGACACCGGCTGTA 1456
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1822 ACCCTGGGCTGTGATCTGCTTCTTCAATATGAGGCTTCCCAATGATACACAGGGTGA 1881
QY 1457 CAATATATCTTGTAGCTGTGATGATCAGTACAGGCTGGAACATTCCTGTGTGTGTGTGTGT 1516
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1882 ATCTACATGTTTGAAGTGTGTGACACCTATGCTGCTCTGTGTGTGTGTGTGTGTGTGT 1941
QY 1517 ATTTTGAATCTGACGAGCGTGTCTGTGATTTATGATGAGATGCTGTGCTAGACAT 1576
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1942 ATTTTGTAGCTTGTGTGATCTCTCTATGTGTAGGCTTGTGAGAGTCTGTGTGAGACATC 2001
QY 1577 GACTGATGTTGGTAAAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1636
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2002 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2061
QY 1637 CCTGCTAATATGACGACTGTGTCTTCTAC 1666
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2062 CCGACCATTTTAAAGCTTATCTTGTCTTC 2091

```

## RESULT 2

```

; Sequence 4, Application US/08655836
; Patent No. 5824486
; GENERAL INFORMATION:
; APPLICANT: Borden, Laurence A.
; APPLICANT: De Vivo, Michael
; APPLICANT: Yokoyama, Midori

```

```

; APPLICANT: Albert, Vivian R.
; TITLE OF INVENTION: GLYCINE TRANSPORTER-TRANSFECTED CELLS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: P.O. Box 5218
; CITY: Princeton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patenting Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/655,836
; FILING DATE: 31-MAY-1996
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Bloom, Allen
; REGISTRATION NUMBER: 29135
; REFERENCE/DOCKET NUMBER: 317743-105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-520-3214
; TELEPHONE: 609-520-3259
; INFORMATION FOR SRO ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2817 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-655-836-4

```

```

Query Match 7.8%; Score 222; DB 1; Length 2817;
Best Local Similarity 54.1%; Pred. No. 1.2e-44;
Matches 503; Conservative 0; Mismatches 415; Indels 12; Gaps 2;

```

```

QY 746 GGAGTCTCGGTGCCCCCATCTGTACTGTGTGTATTCATCGATGGCTCATG 805
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1372 GAATATCTGTGTAGATCAGATGGCCCTTCCGCTTTTCCGGCTGGGTAT 1431
QY 806 GTGTGGAGTCTGCGCCGAGGAGTCAAGATTCGGCAAGCGGCTTACTCTCGCG 865
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1432 GTATATGCAATCGCTGGGAAAGATTAAGACATCAGGAAAGTGTGTCTTACAGCC 1491
QY 866 CTCTCCCATAGCTGTGATGATCATCTTTATTCATCAGCAATCATCTGCGGCTGCT 925
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1492 ACCCTCCCTTATGTCTGTGCTGTGCTGTCTCTCATTCGAGGGGTACCCCTGGAGCT 1551
QY 926 ACTGAGGGAATCTGTCTTCTGTGACGCTCATAGGGGCAAGCTCTGTAGCTGCTGTA 985
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1552 GGAGCCGGTATCTGTGACTTATCATCACTTAAGTGGGAAATCAAGGTCAGAGGTG 1611
QY 986 TGTACTCAGCATCTCAGCAGAGTGTCTTCTGTGACAGTGTGACCGGACCGATCATC 1045
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1612 TGGAGAGATGACACCACTCATGATTTCTTCTGCTGTGCGGCTGGGAGGGCTCATC 1671
QY 1046 ATGTCTCTCTTACAAAGGTTTCAAGATATATATCTACAGGATGCTTGATTTAG 1105
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1672 ACTCTTCTCTTACAAATTCATTAACATGCTGACAGGAGCAAGTTAATTTGAAC 1731
QY 1106 ACTTTGACACCTTTTCAAGTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1165
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1732 TGCACCAACAGTCCCATACATCTTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1791
QY 1166 CTGCGCTACGAATCACTCAGAGTGTGGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1225
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1792 ATGGCCACAGAGGCAAGTCAACATTTGAGAAATGTG--GCTGACCAAGGGCGAGGATTT 1848
QY 1226 GCTTTCATTTCAATACCTGTATGCTGCTGCAAAATTCACCACTAGCTATTTCTGGGTG 1285

```









Db 1345 CTACGGCCCAAGAGGACGATGTTCTCAGCCCTCATCTGTAGCCATGTAAGTGA 1404  
 QY 1433 CTGTTTACTGACACACCGGGTGGACATATATCTTGAGCTTGAATCTACCTAGGGA 1492  
 Db 1405 CTATCTCTCACACCGATGGGGAGTACTGCTGCTCTTCTTGATGACTACAGGCG 1464  
 QY 1493 ACATTCCTGTTGCTTTCTGCGCATTTCTGAACTGCGACGCTGTTCTGATTTATGA 1552  
 Db 1465 AGCTTGCGACTCATTGGGTGGATGATACACATGCTGCTGACACCGGGATACGGC 1524  
 QY 1553 TTGAGCAATCTGTGCTCAGCAATGAGTATGTTGGGTAAAAAGACTGCTTACTGG 1612  
 Db 1525 ATCCAGGGGTTTGTCTGTACATCCATATGCTGCGCTTCAAGCCAGGACTTACTTC 1584  
 QY 1613 CGTCTGCTGGGAGCGTATACCTCTGTAATATGAGAGTGTCTTCTAGCGCTCTT 1672  
 Db 1585 AGGGCGCTGGCTGCTGTTTGTCTCCGCGCACACTTGGCCCTTGGTGTACAGTATC 1644  
 QY 1673 CTGCGCTCTAACAACCTGTGTTCGG 1698  
 Db 1645 GTCAAGTACCAGCCCTCGAATACGG 1670

# RESULT 6 US-07-762-132A-1

; Sequence 1, Application US/07762132A  
 ; Patent No. 5312734

## GENERAL INFORMATION:

; APPLICANT: Uhl, George R  
 ; APPLICANT: Kuhar, Michael J  
 ; APPLICANT: Shimada, Shiochi  
 ; APPLICANT: Kitayama, Shigeo  
 ; APPLICANT: Patel, Amrat  
 ; APPLICANT: Liu, Chien-Liang  
 ; TITLE OF INVENTION: "CDNA Encoding A Dopamine  
 ; TRANSPORTER AND PROTEIN ENCODED THEREBY"  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
 ; STREET: 301 N. Washington St.  
 ; CITY: Falls Church  
 ; STATE: Virginia  
 ; COUNTRY: USA  
 ; ZIP: 22046-0747

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/762,132A  
 ; FILING DATE: 19910920  
 ; CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

; NAME: Murphy, Gerald M.  
 ; REGISTRATION NUMBER: 28,977  
 ; REFERENCE/DOCKET NUMBER: 1173-340P  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 703-241-1300  
 ; TELEFAX: 703-241-0369  
 ; TELEX: 248345

## INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3404 base pairs  
 ; TYPE: NUCLEIC ACID  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Rattus rattus  
 ; IMMEDIATE SOURCE:  
 ; CLONE: pDAT1  
 ; FEATURE:

; NAME/KEY: CDS  
 ; LOCATION: 63..1919  
 ; FEATURE:  
 ; NAME/KEY: polyA\_site  
 ; LOCATION: 3385  
 ; FEATURE:  
 ; NAME/KEY: 5' UTR  
 ; LOCATION: 6..62  
 ; FEATURE:  
 ; NAME/KEY: 3' UTR  
 ; LOCATION: 1920..3384  
 ; US-07-762-132A-1

Query Match 7.34; Score 209.4; DB 1; Length 3404;  
 Best Local Similarity 48.88; Pred. No 1.5e-41;  
 Matches 728; Conservative 0; Mismatches 741; Indels 24; Gaps 5;

QY 239 GAACCGCCAGAACGATATGATGATGTCACACATGTAATCTTGATGCTTCGATCGCT 238  
 Db 228 GAGGCTCAAGAGCGGGGAGACCTGAGCAAGAAATGATTTCTGATACATGATCGGC 287  
 QY 299 ACATCCGCTGTTGGGTAAACGTGTGGCGCTTCCCTTTATCCGCTACAGAAATGAGGA 358  
 Db 288 TTTCGTGTGACCTGGGCAATGTGTGAGGTTTCCCTACCTGTGCTTCAAAAATGTTGA 347  
 QY 359 GGTGCTTCTGTCGATCATGTCATGCTTTTACTTGTGCGCAAGCCTGTGACTAC 418  
 Db 348 GTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 407  
 QY 419 TTAGAGTGTGCTCTGCGCAATTCAGTTCAAGAAACTGTTAAAGTTGCTCAATTTCA 478  
 Db 408 ATGAGAGCTGCTCTGCGCAATTCAGTTCAAGAAAGAGAGAGAGAGAGAGAGAGAGAG 467  
 QY 479 CCGGCCATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 538  
 Db 468 CCGTCTGTAAGAGTGTGGCTTCACTGTTATCTCTCACTCTTCTTCTTCTTCTTCTTCT 527  
 QY 539 TAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 598  
 Db 528 TACAATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 587  
 QY 599 CCATGGCTATTTGTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 650  
 Db 588 CCATGGATCCATGCAACCAACACCTGGAATAGCCCAACTGCTGATGATGATGATGATGAT 647  
 QY 651 ----CACTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 706  
 Db 648 AACTTACGAGAGCGCTGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 707  
 QY 707 TTTTGAAGAACGTTCTCAACAAAGCATGATGATGATGATGATGATGATGATGATGATGAT 763  
 Db 708 TTTGAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 767  
 QY 764 ATCTGTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 823  
 Db 768 CGGTGCAAGCTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 827  
 QY 824 CGAGAGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 883  
 Db 828 AAGGAGTAAGACCTCAG 887  
 QY 884 ATGATCACTTATTCATCAGCAATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 943  
 Db 888 CTCACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 947  
 QY 944 TTGTCACAGCTTAAAGGCGAAACTCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1003  
 Db 948 TACCTAGTGTGATCTTCAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1007  
 QY 1004 CAAGTGTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1063  
 Db 1008 CAGGTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1067

QY 1064 GCTTCAGACATTAATCTACAGGAGATGCTGGATTTGACATTTGGACACCTTTACA 1123  
 DB 1068 AATATCACCATACTAGTATAGAGCAATCACCACCTCATTTACACCTCCGAGACA 1127  
 QY 1124 AGTTCTGTCTGGGTGACGATCTTCGGTATCTTGGTAACCTGGCTGACAACTCAAC 1183  
 DB 1128 ACCTTCCTCTGGGTGCTGCTGCTCTTCCTTCCTGGGGTAAATGGCACAGAGCAAT 1187  
 QY 1184 TCAGAGTGGAGATGCTGGTCTGGCTGACAGCTTCCTTCCTTCATTAACCT 1243  
 DB 1188 GTGCCCATCAGAGATGTGCC---ACAGATGACCTGGGCTATCTTCACTATCTT 1244  
 QY 1244 GATGCCATTCGCAAAACATTCACACCTAGCTATTTCTGGTCTCTTCTCTGANG 1303  
 DB 1245 GAGGCGATCGCCACACTCCCGCTGCTTCTGGCTGGGCTGCTGCTCTCTCTCAATG 1304  
 QY 1304 TCGGTGCTGGGTATGGGCTCATCCGCTGCTGCTGATGACCTTTCAACCACTGGCGAT 1363  
 DB 1305 CTCACCTGCGGTATGACAGATGCAATGGGGGACATGAGTACATCTGGGCTGCTC 1364  
 QY 1364 GACGCGTTCGCCACGCTGAC---CCACGCTACATGTCAGGATGACCTGTTCTGGG- 1418  
 DB 1365 GATGAGTTCACAGCTGCTACATGAGGATGAGAGCTCTTCACTTGGCATTTGCTGGCT 1424  
 QY 1419 -GTTTCTGCTTGGACTGTTTACTGCAACCCGCTGACAAATATTTCTGAGCTTGA 1477  
 DB 1425 ACTTCTGCTGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1484  
 QY 1478 GATCCTACGCGTGAACATTCCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1537  
 DB 1485 GACCATTTTGCAGCTGCGACATCTATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1544  
 QY 1538 TTCTGGATTTATGATGATGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1597  
 DB 1545 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1604  
 QY 1598 ACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1657  
 DB 1605 CCCAAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1664  
 QY 1658 TTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1710  
 DB 1665 GTCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1717

RESULT 7  
 US-08-301-722A-1  
 : Sequence 1, Application us/08301722A  
 : Patent No. 5756307  
 :  
 : GENERAL INFORMATION:  
 : APPLICANT: UNL, George R.  
 : APPLICANT: Vandenberg, David  
 : APPLICANT: Persico, Antonio  
 : TITLE OF INVENTION: SEQUENCE OF HUMAN DOPAMINE TRANSPORTER  
 : TITLE OF INVENTION: CDNA  
 : NUMBER OF SEQUENCES: 10  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP  
 : STREET: P.O. Box 747  
 : CITY: Falls Church  
 : STATE: Virginia  
 : COUNTRY: USA  
 : ZIP: 22040-3487  
 :  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/301,722A  
 : FILING DATE: 07-SEP-1994  
 : CLASSIFICATION: 435  
 : ATTORNEY/AGENT INFORMATION:

: NAME: Murphy Jr., Gerald M.  
 : REGISTRATION NUMBER: 28,977  
 : REFERENCE/DOCKET NUMBER: 1173-406P  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 703-205-8000  
 : TELEFAX: 703-205-8050  
 : INFORMATION FOR SEQ ID NO: 1:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 3919 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: double  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: CDNA  
 : HYPOTHETICAL: NO  
 : ORIGINAL SOURCE:  
 : ORGANISM: Homo sapiens  
 : TISSUE TYPE: Brainstem  
 :  
 : FEATURE:  
 : NAME/KEY: CDS  
 : LOCATION: 102..1961  
 : OTHER INFORMATION: /function="dopamine transport"  
 : OTHER INFORMATION: /product="HUMAN polypeptide, see Fig. 5 (Hdalt)  
 :  
 : FEATURE:  
 : NAME/KEY: misc\_RNA  
 : LOCATION: 2724..3117  
 : OTHER INFORMATION: /function="unknown"  
 : OTHER INFORMATION: /label="VNTN\_region  
 :  
 : US-08-301-722A-1  
 :  
 : Query Match 7.2%; Score 204.8; DB 1; Length 3919;  
 : Best Local Similarity 48.6%; Pred. No. 2.1e-40;  
 : Matches 727; Conservative 0; Mismatches 742; Indels 27; Gaps 5;

QY 239 GAACGCCAGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 298  
 DB 267 GAGGCCAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATG 326  
 QY 299 ACATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 358  
 DB 327 TTTGCTGAGACCTGCGACAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 386  
 QY 359 GGTGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 418  
 DB 387 GGTGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 446  
 QY 419 TTAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 478  
 DB 447 ATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 506  
 QY 479 CCGGCCATGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 538  
 DB 507 CCCATAGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 566  
 QY 539 TACGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598  
 DB 567 TACAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 626  
 QY 599 CCATGGGCTATTTGACAGCTGAG 658  
 DB 627 CCTGAGATCAGCTGACCAACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 686  
 QY 659 G-----CATAGTCAACAACATCAACAACATCAACAACATCAACAACATCAACAACATCA 703  
 DB 687 GACTCAGTGAAG 746  
 QY 704 TACTTTTGAAG 760  
 DB 747 TACTTTGAAG 806  
 QY 761 CCCATCGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 820  
 DB 807 CCGCGGTGAGAGCTCAG 866

QY 821 GCCGAGAGAGTCAAGAGTTCGGGCAAGCGCTACTTCCGCGCTTCCCATAGTT 880  
Db 867 TGGAGAGGCGTGAAGACCTCAGGGAAGGTGTATGATTCAGACGACCAATGCCATAGCTG 926  
QY 881 GTATATGATCACTTATTCATCACCACATCATCTGCGCGGTCTACTAGCGCATCTG 940  
Db 927 GTCTCACTGCTCCCTGCTCCGTGGGGTCCACCTCTCTGAGCCATGAGCGCATCAGA 986  
QY 941 TTCTGTGTCAGCGCTCAATGGGGAAACCTTGGAGTGGGTATGATGCTACTGACAGTC 1000  
Db 987 GCATACCTGAGCGTGTACTCTTCAACCGGCTTGCAGGCGCTGTGTGGATTGACGCGCC 1046  
QY 1001 ACGCAAGTGTCTTCTCTCTGACAGTGTGACCGGACCGATCATCATGTTCTCTTAC 1060  
Db 1047 ACCAGAGTGTCTTCTCTGAGGCGTGGGGTGTGTATGCGCTTCTCAGAGTAC 1106  
QY 1061 AAGGTTTCAGACATATATCTACAGGATGCTTGGATTGTTCAGACTTGGACACTT 1120  
Db 1107 AACAGTTTCACCAACATCTCTACAGGACGCGATGTCTACACACTTCATCACTCCCTG 1166  
QY 1121 ACAAGTTTCTGTGGGTGACAGATCTTCGATATCTGTGAACCTCGGCTAGCAATC 1180  
Db 1167 ACAGAGTCTCTCTCCGCTTCCGCTTCTCTCTCTCTGAGGATGATGACAGAGAC 1226  
QY 1181 AACTCAGAGGTGGAGATGTGTGCGGTGCGGCTACAGTCTTGTCTTCAATTCATAC 1240  
Db 1227 AGTGTCCCATCGGGGAGGTGGCC---AAGGACGGGCGAGGGGTGATCTTCAATCATC 1283  
QY 1241 CCGTATGCCATTTGCCAAAATTCACACCTCAGCTATCTCGGTGCTGTCTTCTCTATG 1300  
Db 1284 CCGGAGGCCATTCGCGAGCTCCCTCTCTCAAGCGGCGCGGTCTTCTTCAATCATG 1343  
QY 1301 ATGTGCGGTGCTGGGTATCGGCTCATCCGCTCTGTGCTATGACTTTCACACATTTGGC 1360  
Db 1344 CTGTCTACCTGTGGTATCGAGAGGCAATGGGTGTGTGATGATGATGATCAGCGGCTC 1403  
QY 1361 ATGAGCGGTCCAC---GTGTACCCACGCTCTACATGTCAAGCATGACTGTCTTG 1416  
Db 1404 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1463  
QY 1417 CGGT---TTCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1474  
Db 1464 GCGAGCT 1523  
QY 1475 GTAGTACATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1534  
Db 1524 CTGAGACATTTTGAAGCGGAGCGTCCATCTCTGTGATGATGATGATGATGATGATG 1583  
QY 1535 GTGTCTGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1594  
Db 1584 GTGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1643  
QY 1595 AAGAGTGTGCTTACTGCGCTCTGTGCGGCGGTATGATGATGATGATGATGATG 1654  
Db 1644 CGGCGCAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1703  
QY 1655 GTGTCTTCTACGCT 1710  
Db 1704 GTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1759

RESULT 8  
US-09-191-468-121  
; Sequence 121, Application US/09191468A  
; Patent No. 6416975  
; GENERAL INFORMATION:  
; APPLICANT: Gallager, Michael J.  
; APPLICANT: Burgess, Loyd R.  
; APPLICANT: Brunden, Kurt R.  
; TITLE OF INVENTION: Human Glycine Transporter Type 2  
; FILE REFERENCE: 1231US01  
; CURRENT APPLICATION NUMBER: US/09/191,468A  
; CURRENT FILING DATE: 1998-11-12

QY 746 GGAGGTCTGGTCCCGCATCTGATGATGATGATGATGATGATGATGATGATGATG 805  
Db 1159 GAATATCTGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1218  
QY 806 GTGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 865  
Db 1219 GTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1278  
QY 866 CTCTCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 925  
Db 1279 AGCTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1338  
QY 926 ACTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 985  
Db 1339 GGAAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1398  
QY 986 TGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1045  
Db 1399 TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1458  
QY 1046 ATGTCTCTCTTACCAAGGTTTCAAGATATATCTACAGGATGATGATGATGATG 1105  
Db 1459 ACTCTCTCTTACCAAGATATATCTACAGGATGATGATGATGATGATGATGATG 1518  
QY 1106 ACTTGGACCTTTTACCAAGATATATCTACAGGATGATGATGATGATGATGATG 1165  
Db 1519 TGCACCAAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1578  
QY 1166 CTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1225  
Db 1579 ATGGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1635  
QY 1226 GCTTTCATTTTACCAAGTGGATGATGATGATGATGATGATGATGATGATGATG 1285  
Db 1636 GCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1695  
QY 1286 CTGTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1336  
Db 1696 ATCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1755  
QY 1337 CTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1396  
Db 1756 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1815  
QY 1397 TCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1456  
Db 1816 ACTCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1875  
QY 1457 CATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1516  
Db 1876 ATTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1935  
QY 1517 ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1576  
Db 1936 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1995  
QY 1577 GAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1636





```
QY 1493 ACATTCCTTGTCTTCTGCGCATTTCTGAATCCGAGCGGTGTTGGATTATGGA 1552
      |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1607 GGCCGACAGTGTGTCACCGCGGCGGCTCATGAGAGCCCTCCGCGTCTTGGTTATGGA 1666
QY 1553 TTGGAGATCTGTGCTAGACATTTGATGTTGGGTAAAGAGTGTGCTTACTGG 1612
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1667 ATACTGAGTTCTGACAGCGATGTGAAGAGATGCTGGGCTTCAACCGGATGTTTGG 1726
QY 1613 CGTCTGCTGGGGCGGTAATCACTCCT 1639
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1727 AGGATCTGCTGGGCGCATCAAGCCCT 1753

RESULT 11
US-09-191-468-119
; Sequence 119, Application US/09191468A
; Patent No. 6416975
; GENERAL INFORMATION:
; APPLICANT: Gallagher, Michael J.
; APPLICANT: Burgess, Lloyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 12311US01
; CURRENT APPLICATION NUMBER: US/09/191,468A
; CURRENT FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 119
; LENGTH: 2394
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2391)
; OTHER INFORMATION: Seq 49 HGLYT2; nt 1-2394
US-09-191-468-119

Query Match 6.8%; Score 194.8; DB 4; Length 2394;
Best Local Similarity 52.3%; Pred. No. 4,7e-38;
Matches 486; Conservative 0; Mismatches 432; Indels 12; Gaps 2;

QY 746 GGAAGTCTGCGTCCCATCTGCTACTGTTGTTGTTATTCATCGATGAGCTCATG 805
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1159 GAATATCTGCGGAGATGAGTGGCCACTAGCTCTGCTCTTCCGCTGGGCTGATCATT 1218
QY 806 GTGTTCGAGATCTGCGGCCAGGAGTCAAGATTCGCGCAAGCGGCTTCTCTCGCG 865
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1219 GTGTATCATCTGTGGCTAAAGATCAAGACTTCAGAAAGTGTGTACTTACCGGCC 1278
QY 866 CTCTCCCATACGTTGTCATGATCACTTTATTCATCCACCAATCATCTGCCGGTCT 925
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1279 ACGTTCGGTATGCGGCTAGTATGATCTCTCTCTCTCGAGAGATCACTGCTGGAGCT 1338
QY 926 ACTGAGCGCATCTGTTCTTCTGTCACGCTCATATGGCGAAATCTTGAAGCTTGGTGA 985
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1339 GGAAGCTGGATCTGTACTTATCATCACCAAGTGGGAGAAACTCAGGATGCCAGGTG 1398
QY 986 TGGATCACAGTCAAGCAAGTGTCTCTCTGACAGTGTGTACCGGACCGCATCATC 1045
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1399 TGGAAAGATCTGCGCATCTAGATTTCTTATCTGTGATGAGGAGGCGCTGATC 1458
QY 1046 ATGTTCTCTTACACAGGTTTCAGACATATATATACAGAGGATGTTGGATTGTTAG 1105
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1459 ACGTCTCTTCTTACAAATATCCACACAGCTGTACAGGACATCTTAATTTGTACAC 1518
QY 1106 ACTTTGACACCTTTACAGTCTTGTCTGTGTCAGCATCTTGGTATCTTGTGTAAC 1165
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1519 TGCACCAACAGTGCACAGATCTTTGCGGCTTCCTCACTCTCCGTATGCGCTTC 1578
QY 1166 CTGCGGTACAGTCAAGCTCAAGAGGTGGAGATGTGTGCTGTGCGGTACAGATCTT 1225
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1579 ATGGCCATGACGCAAGTCAACATGTAGAGATGTG--GCTGACCAAGGCGCATTT 1635
```

```
QY 1226 GCTTCAATTCATACCCGATGATGCGATTCGCAAAACATTCACACTGATCTCGGTG 1285
      |||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1636 GCATTTGTGTTTACCGGAGGACTTAACGAGGCTGCTCTCTCTCGGTTCTGGGCATC 1695
QY 1286 CTGTCTTCTTCGATGATGTGCGGTGCGG-----TATCGGCTCATCGTGTCTGTG 1336
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1696 ATCTTTTCCGATGCTCTCTACCTTGTGACTGTACACTATGTTGTGCTCATGAGACCC 1755
QY 1337 CTATGACTTTCACACATTTGGCGATGAGCGGTTCCACGCTGATCCACCGCTTACATG 1396
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1756 ATACTGACCTTCACATCTAGACAGAGATTTCCAAATGCTTACGACACAGCAGTGT 1815
QY 1397 TCAGCATGACCTGCTCTGCGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1456
      |||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1816 ACTCTGGCGCTGCTCATTTGTTTCTTATCATGAGGTTTCCAAATGATCACTGAGGTGA 1875
QY 1457 CAATATATCTTGTGAGCTTGTAGATCACTAGCGTGAACATTCCTTGTGCTTTCTGCGCC 1516
      |||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1876 ATTTACATGTTTACGCTTGTGACACCTATGCTGCTCTCTCTCTCTCTCTCTCTCTCT 1935
QY 1517 ATTTCTGACCTGCGAGCGGTCTCTGATTTATGATTTGAGATCTGTGCTTGTGATCATT 1576
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1936 ATTTTGTAGCTCGTGGGATCTCTTATGTATGTGCTGCAAAAGATCTGTGAAGATATA 1995
QY 1577 GAGTTCAATGTTGGTAAAGAGTGTGCTTACTGCGGCTCTCTGCGGCGTAAATCACT 1636
      |||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1996 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2055
QY 1637 CTGTCTATATGACGAGCTGTGTTCTTAC 1666
      |||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2056 CCAACCATTTTAACTTATCTTCTTCTTCT 2085
```

```
RESULT 12
US-09-191-468-123
; Sequence 123, Application US/09191468A
; Patent No. 6416975
; GENERAL INFORMATION:
; APPLICANT: Gallagher, Michael J.
; APPLICANT: Burgess, Lloyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 12311US01
; CURRENT APPLICATION NUMBER: US/09/191,468A
; CURRENT FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 123
; LENGTH: 2394
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2391)
; OTHER INFORMATION: SEQ ID NO:26 [MO98/07854 (PCT/US97/14637)] Allele1x
; OTHER INFORMATION: Sequence: nt 1-2394; nt 304 may be G; nt 371 may
; OTHER INFORMATION: be T; nt 836 may be A; nt 1116 may be G; nt 1831
; OTHER INFORMATION: may be G; nt 2382 may be A or T; nt 2385 may be G;
; FEATURE:
; OTHER INFORMATION: nt 2388 may be A
US-09-191-468-123
```

```
Query Match 6.8%; Score 193.2; DB 4; Length 2394;
Best Local Similarity 52.2%; Pred. No. 1.2e-37;
Matches 485; Conservative 0; Mismatches 433; Indels 12; Gaps 2;

QY 746 GGAAGTCTGCGTCCCATCTGCTACTGTTGTTGTTATTCATCGATGAGCTCATG 805
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1159 GAATATCTGCGGAGATGAGTGGCCACTAGCTCTTCCGCTTCCGCTTGGGCTCATTT 1218
QY 806 GTGTTCGAGTCTGCGGCCAGGAGTCAAGATTCGCGCAAGGCGGCTTCTTCTCGCG 865
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

Db	1219	GTGATGATGCGTTGGCGTAAAGAACTAAGAACTTCAGGAAAGTGGTACTTTCAGCGCC	1278
Qy	866	CTTTTCCCATACGTTTGTCATGATCATCTTTATTCATCACCAATATCATCTCCCGGCT	925
Db	1279	ACGTTCCCGGTATGTCGTACTCGTGTATCCCTCTCATCCAGGAGATCACCTCTCCGTGACCT	1338
Qy	926	ACTGACGGCATCTGTTTGTGTGTGACGGCTCAATGGGGGAAACCTCTTGACCTGGGNTA	985
Db	1339	GGAGCTGGGATCTGTGTACTTATCATCACCCAAAGTGGGAAATCTACGAGATGCCAGGTG	1398
Qy	986	TGTATCTACGAGTACGCAAGTGTCTTCTCTGTGACAGTGTGCACCGGACGATCATC	1045
Db	1399	TGGAAAGATGCTGCACATCAATTTTCTTCTTATCTGTGATGAGGGAGGCGCTGATC	1458
Qy	1046	ATGTTCTCTCTTACAAAGGTTTTCAGACATATATCTACAGGAGATGCTTGGATTTGTACG	1105
Db	1459	ACTCTCTTTCTTACAAATAATCTCACAACTGCTACAGGACACTCTATATGTCTACC	1518
Qy	1106	ACTTTGACACCTTTACAAAGTTTCTTGTGCTGGGTGACAGATCTTCGATCTCTGTATAC	1165
Db	1519	TGCACCAACAGTGGCACAAAGCATTTTGTGCGGCTTCGTATCTTCTCGTTATGTGGCTTC	1578
Qy	1166	CTCGCGTACGAACTCACTACAGAGTGGAGATGTTGTCGTGCTGGCGGTACAGTCTT	1225
Db	1579	ATGGCGCAATGAGCAGCAAGTCAATGATGAGATGTG---GCAGACAAAGGCGCAGGACTT	1635
Qy	1226	GCTTTCATTTCAATACCCGATGATGACCATTCGCAAAACATTCAAACCTACAGCTATTCGCGTG	1285
Db	1636	GCAATTTTGCGTTTACCGGGAAGCCTTAAACCAAGCTGCTCTCTCTCCGTTCTGGGCCATTC	1695
Qy	1286	CTGTTCTTCTGATGATGTGCGTCTGGGATTCGAGCTCATCCGTCGCTCTCATATGCACT	1345
Db	1696	ATCTTTTTTCCGATGTCCTCACTCTTGATGACATATGTTTGGCACCATGTGAGAC	1755
Qy	1346	TTCAACACATTTGGGGATGAGCGCTTCCAGCTGTACCCAGCCGTCATATGTACAGCATG	1405
Db	1756	ATAGTAGCTCCATCTCAGACGAGTTTCCAAATACATCAGCAGACACAAACCAAGTGT	1815
Qy	1406	AC-----CTGTTCTTCTGCGGTTTCCCTGCTGGAGACTTGTTACTGCACACCGGAGGA	1456
Db	1816	ACTCTGGGCTCTGCAATTTGTTTCTTCAATGAGGTTTCCAAATGATACACAGGATGA	1875
Qy	1457	CAATATATCTTGAAGCTTGTGATCAGTACAGCGTGAACATTCCTTGTCCTTTCTGCGCC	1516
Db	1876	ATTTCATGTTTCAAGCTTGTGGACACATATCTCCCTCTATGCCCCCTGTATATGCCC	1935
Qy	1517	ATTTCGTAACTCGAGCGGCTGTCTGAGTTTATGAGTATGGAGAAATCTGTGCTGTAGACTT	1576
Db	1936	ATTTTTGAAGCTCTGGGGATCTCTTATGTGTATGTGCTTGCAAAGATTTCTGTGAAGATATA	1995
Qy	1577	GAGTTCATGTTGGGTAATAAGACTGTGTACTGTGCGCTCTCTCTGGGCGGTATCACT	1636
Db	1996	GAGATGATGATGGATTCCAGCCCTAACATCTTCTGGAAGAACTCTGTGGCATTTGTAAAC	2055
Qy	1637	CTGTGATTAATGACGAGCTGTCTCTTAC	1666
Db	2056	CCAACCATTTTAACTTATCTTTTCTTC	2085

RESULT 13  
US-08-700-013B-18  
Sequence 18, Application US/08700013B  
Patent No. 5919653  
GENERAL INFORMATION:  
APPLICANT: Albert, Vivian R.  
APPLICANT: Kowalski, Leslie R.Z.  
APPLICANT: Borden, Laurence A.  
APPLICANT: McKelvy, Jeffrey F.  
TITLE OF INVENTION: Human Glycine Transporter  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dechert Price & Rhoads  
STREET: 997 Lenox Drive, Building 3, Suite 210

CITY: Lawrenceville  
STATE: NJ  
COUNTRY: USA  
ZIP: 08543  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/700.013B  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Bloom, Allen  
REGISTRATION NUMBER: 29,135  
REFERENCE/DOCKET NUMBER: 317743-108  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 609-520-3214  
TELEFAX: 609-520-3259  
TELEX:  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2397 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

Query Match	Score	DB 2;	Length
6.88;	193.2;		2397

Matches 485; Conservative 0; Mismatches 433; Indels 12; Gaps 2;

746 GGAGGTCGCGTGCCCCCATCTGGTACTTGGTGTGCTATTCAATCGCATGGCTCATG 805

Db 1159 GAATATCCTGGCGAGATCAGGTGGC<sup>•</sup>ACTAGCTCTCTGCCCTCTCCCTGGCTGGGTCATT 1218

806 GTGTCGGAGTCGTCGCCCCGAGGAGTCAAGAGTTCGCGCAAGCGGCCTACTTCCTCGCG 865

Db 1219 GTGATGCAATCGTTGCTAAAGGAATCAAGACTTCAGGAAGAGTGGTGTACTTCACGGCC 1278

**098**

[illegible][illegible]

20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

DB 1359 GGAGCTGGGATCTGGTACTTCATCACACCCAAATGGGAGAAATCTACGGATGCCACGGTG 1396

986 TGGTACTCAGCAGTCACGCAAGTGTCTCTCTCTGACAGTGTGCACCGGACCGATCATC 1043

Db 1399 TGGAAAGATGCTGCCACTCAGATTTCCTCTTTATCTGCTGCATGGGAGGCCCTGATC 1458

1046 ATGTTCTCCTTACACGGTTTCAGACATATATCTACAGGATGCTTGGATTGTTACG 1105

Db 1459 ACTCTCTCTTCTTACACAATTCACACAACCTGCTACAGGACACTCTAATTGTCACC 1518

1106 ACTTTGACACCTTACAAAGTTCTTGTCTGGGTGCACGATCTTCGGTATCCTTGGTAAC 1165

Db 1519 TGCACCAACAGTGCACACAGCATCTTTGCCGCGTTCGTCAATCTCTCCGTTATCGGCTTC 1578

1166 CTCGCGTACGAAGTCACTCAGAGGTGGAGATGTGGTCCGTCCTGGCGGTACCACTCTT 1225

[illegible][illegible][illegible]

UU 1030 GCAIIIGIGGIIATCCCGAAGCCIIATCCAGGCIAGCIIICIIICCCGIIICGGCCATC 1031

1286 CTGTTCTTCCGTGATGATGTCGGTGTCTGGGTATCGGCTCATCCGTGGCTCTGTCTATCGACT 1343



Db 1696 ATCTTTTCTCTATGCTCTCACTCTTGAGCTTGACACTATGTTGCCACATCGAGACC 1755  
 QY 1346 TTCACACATTTGGCCATGAGCGCTTGCCACGATACCCAGCGTCTACATGTCAGGATG 1405  
 Db 1756 ATAGAGACCTCATCTCAGACAGATTTCCCAAGTACCTAGCACACACAGGAGTGT 1815  
 QY 1406 AC-----CTGTCCTGCGGTTCCCTGCTTGAGCTTGTTACTGCACACCGGGTGA 1456  
 Db 1816 ACTCTGGCGTCTGATTTGTTCTTCATCATGAGGTTTCCAATGATCTCAGGTTGA 1875  
 QY 1457 CAATATTTCTGAGCTTGTATATCATACGCTGGAACATTCCTTGCTTTCTGCGGC 1516  
 Db 1876 ATTTCAGTTCTACCTTGTGACACATATGCTGCTCTATGCCCTTGTCATCATGCGC 1935  
 QY 1517 ATTTCGACCTGCGAGCTGTTCTGATTTATGATTTGAGATTTGAGATTCGCTAGACAT 1576  
 Db 1936 ATTTTGAAGCTGCTGGGATCTCTTATGATGCTGCAAGATTCGTGAAGATATA 1995  
 QY 1577 GAGTTCATGTTGGTAAAGACCTGCTGCTTACTGCGCTCTGCTGGGCGTAAATCACT 1636  
 Db 1996 GAGATGATGATTTGATTCAGCCATCAATCTTGGAAGTCTGCTGGGCAATTTGTAAC 2055  
 QY 1637 CCTGCTATATGAGACTGTGTTCTTAC 1666  
 Db 2056 CCAACCATTTTACCTTATCTTGCTTC 2085

## RESULT 14

US-08-700-013B-20

; Sequence 20, Application US/08700013B  
 ; Patent No. 5919653

## GENERAL INFORMATION:

; APPLICANT: Albert, Vivian R.  
 ; APPLICANT: Kowalski, Leslie R.Z.  
 ; APPLICANT: Borden, Laurence A.  
 ; APPLICANT: McElvly, Jeffrey F.  
 ; TITLE OF INVENTION: Human Glycine Transporter  
 ; NUMBER OF SEQUENCES: 41  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: Dechert Price & Rhoads  
 ; STREET: 997 Lenox Drive, Building 3, Suite 210  
 ; CITY: Lawrenceville  
 ; STATE: NJ  
 ; COUNTRY: USA

ZIP: 08543

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/700.013B

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:

NAME: Bloom, Allen

REGISTRATION NUMBER: 29,135

REFERENCE/DOCKET NUMBER: 31743-108

TELECOMMUNICATION INFORMATION:

TELEPHONE: 609-520-3214

TELEFAX: 609-520-3259

TELEX:

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 2397 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-700-013B-20

Query Match 6.8%; Score 193.2; DB 2; Length 2397;  
 Best Local Similarity 52.2%; Pred. No. 1.2e-37;  
 Matches 485; Conservative 0; Mismatches 433; Indels 12; Gaps 2;

QY 746 GAGAGCTGCGGCCCCCATCTGATCTGATGCTGCTGCTATTCATGCAATGGCTATG 805  
 Db 1159 GAATATCTCTGGGAGATATAGGAGGACATAGCTCTCTCTCTCTCTCTCTCTCTCTCT 1218  
 QY 806 GTTTCGAGATGCTGCGCCGAGAGTTCGAGATTCGAGATTCGAGATTCGAGATTCGAG 865  
 Db 1219 GTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1278  
 QY 866 CTCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 925  
 Db 1279 ACGTTCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1338  
 QY 926 ACTGACGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 985  
 Db 1339 GAGCTGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1398  
 QY 986 TGTACTGACGATCAGCAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1045  
 Db 1399 TGGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1458  
 QY 1046 ATGTTCTCTCTTCAACGCTTCTGATGATGATGATGATGATGATGATGATGATGATG 1105  
 Db 1459 ACTCTCTCTCTTCAACGCTTCTGATGATGATGATGATGATGATGATGATGATGATG 1518  
 QY 1106 ACTTGGACACCTTACAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1165  
 Db 1519 TGCACCAACAGTGCACAGCATCTTCCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTG 1578  
 QY 1166 CTCGCTGACGATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1225  
 Db 1579 ATGGCAATGACCAAGTCAATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1635  
 QY 1226 GCTTTCATTTCAACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1285  
 Db 1636 GCATTTTGCTTACCGGAGACCTTACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1695  
 QY 1286 CTGTTCTCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1345  
 Db 1696 ATCTTTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1755  
 QY 1346 TTCACACATTTGGGAGAGAGCGCTTCCACAGTATACCCAGCTCTCATGATGATGATG 1405  
 Db 1756 ATAGTACCTTCATCTCAGACGAGTTCCTCAAGTACCTAGCACACACAGCCAGTGT 1815  
 QY 1406 AC-----CTGTCCTGCGGTTCCCTGCTTGAGCTTGTTACTGCACACCGGGTGA 1456  
 Db 1816 ACTCTGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1875  
 QY 1457 CAATATTTCTGAGCTTGTATATCATACGCTGGAACATTCCTTGCTTTCTGCGGC 1516  
 Db 1876 ATTTCAGTTCTACCTTGTGACACATATGCTGCTCTATGCCCTTGTCATCATGCGC 1935  
 QY 1517 ATTTCGACCTGCGAGCTGTTCTGATTTATGATTTGAGATTTGAGATTCGCTAGACAT 1576  
 Db 1936 ATTTTGAAGCTGCTGGGATCTCTTATGATGCTGCAAGATTCGTGAAGATATA 1995  
 QY 1577 GAGTTCATGTTGGTAAAGACCTGCTGCTTACTGCGCTCTGCTGGGCGTAAATCACT 1636  
 Db 1996 GAGATGATGATTTGATTCAGCCATCAATCTTGGAAGTCTGCTGGGCAATTTGTAAC 2055  
 QY 1637 CCTGCTATATGAGACTGTGTTCTTAC 1666  
 Db 2056 CCAACCATTTTACCTTATCTTGCTTC 2085

## RESULT 15

US-09-182-728A-1

; Sequence 1, Application US/09182728A

```

Patent No. 6238883
GENERAL INFORMATION:
APPLICANT: BROWN, ANTHONY
APPLICANT: CHAPMAN, CONRAD GERALD
APPLICANT: GLOGER, ISRAEL SIMON
APPLICANT: EVANS, JOANNE RACHEL
APPLICANT: CAIRNS, WILLIAM
APPLICANT: HERDON, HUGH
TITLE OR INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GP-30176
CURRENT APPLICATION NUMBER: US/09/182.728A
CURRENT FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: 9818890.7
PRIOR FILING DATE: 1998-08-28
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2863
TYPE: DNA
ORGANISM: HOMO SAPIENS
US-09-182-728A-1

```

```

Query Match      6.7% Score 191.6; DB 4; Length 2863;
Best Local Similarity 52.0%; Pred. No. 3,1e-37;
Matches 484; Conservative 0; Mismatches 434; Indels 12; Gaps 2;

```

```

QY 746 GGAGCTCTCGGTGCCCCCATCTGCTACTGTGCTGTCTATTCATCGCATGGCTCATG 805
DB 1414 GAATATCTCTGCGGAGTACAGTGGCCACTAGCTCTGCTGCTCTGCTGCTGCTGCTATT 1473
QY 806 GTGTTCGAGTCTGCGCCGAGAGTCAAGTTCGCGAAAGCGGCTTACTTCTCTCG 865
DB 1474 GTGTATGATCGTGTGGCTTAAGAAATCAAGACTTCAGAAAGTGTGTACTTACGGCC 1533
QY 866 CTCTCCCATACGTTGATGATCACTTTATTCATCACACATCATCTGCGCGTCT 925
DB 1534 ACGTCCCGATGCTGCTACTGCTGATCTCTCTCATCCGAGAGTCACTGCTGCGAGCT 1593
QY 926 ACTGAGCGCATCCGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 985
DB 1594 GGAGCTGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1653
QY 986 TGTACTACAGCATCAGCAGCAAGTGTCTCTCTGACAGTGTGACCGGACCGCATCATC 1045
DB 1654 TGGAAAGATGCTGCGACTCAGATTTCTCTCTTATCTGCTGATGGGAGGCGCTGATC 1713
QY 1046 ATGTTCTCTCTCAACAGGTTTACAGATATATCTACAGGATGCTTGGATTGTACG 1105
DB 1714 ACTCTCTCTCTTACACAAATTCACAACTGCTACAGGACACTCTAATGTGACCC 1773
QY 1106 ACTTTGGACACCTTTACAAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1165
DB 1774 TGCACACACAGTGCACAGCATCTTGGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1833
QY 1166 CTCGCGTGAACACTCACTCAGAGTGGAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1225
DB 1834 ATGGCATATGAACGAAAGTCAACATGAGATG--GCAGACCAAGGCGCAGGCATT 1890
QY 1226 GCTTCAATTCAACCTGATGACCAATGCCCCAAACATTCACACTCACTATTCTCGGTG 1285
DB 1891 GCATTTGTGTTTACCCGGAAGCCTTAACAGGCTGCTCTCTCTGCTTCTGCGGCATT 1950
QY 1286 CTGTTCTCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1345
DB 1951 ATCTTTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2010
QY 1346 TTCAACACATGGGAGTGAAGCGTTCACAGTATCCACACGCTTACATGTCAGCATG 1405
DB 2011 ATAGTACCTCACTGACAGAGATTTCCAAAGTACCTACAGCACACAAAGCAGATGTTT 2070
QY 1406 AC-----CTGTTCTGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1456
DB 2071 ACTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2130

```

```

QY 1457 CAATATATCTGAGCTGTAGATCACTACAGGAGGAAACATCTGTGCTTCTGCGCC 1516
DB 2131 ATTACATCTTCAAGCTTGTGAGACCTTATGCTGCTCTATGCCCTTATCATCATATGCC 2190
QY 1517 ATTCTGAATCGCAGCGCTGTCTGATTTATGATTTGAGAAATCTGTGCTAGACATT 1576
DB 2191 ATTTTGAAGCTGTGGGATCTCTTATGATGCTTCAAAAGATTTCTGTGAAGATATA 2250
QY 1577 GAGTTCATGTGGGTAAAAAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1636
DB 2251 GAGATGATGATTTGATTCAGGCTTACATCTTCTGAAAGTCTGCTGCGCATTTGTAACC 2310
QY 1637 CCTGTATATGACGACTGTTCTCTTAC 1666
DB 2311 CCACCATTTTAACTTTATCTTCTGCTTC 2340

```

Search completed: July 26, 2003, 03:09:31  
Job time : 142 secs



Db 534 CCCATCTGAAAGGTGGGCTTCACGGTCACTCTCTACATGATGTCGCTCTTC 593  
QY 539 TACGTGGTATCTGTGTCTCTGTCTATTAATAGCTATGATCCAGGCACCTT 598  
Db 594 TAAACGCTATCTATCCGCTGGGCTGACCTATCTCTCTCTCTCTCTCTCT 653  
QY 599 CCAATGGGCTATTTGTGACCTGATGGGAGACCTGCTACCTCAGATCCAGACTTGT 658  
Db 654 CCTGTGATCAGTCAACCACTCTGGAAACAGCCCAACCTGCGATGCGCATCTG 713  
QY 659 G-----CATAGTCAACCACTACCATGATGATCCAGCACTGCTCAATC 703  
Db 714 GACTCCAGTGGAGACAGCTCGGCTCCACAGACATTTTGGAGCAACCTGCTG 773  
QY 704 TACTTTTGGAAACAGTCTCCAAACAAAGGATGGAATG---AAGGAGGCTTCG 760  
Db 774 TACTTTGAAAGTGGCTGCTGACCTCCACCAAGAGCCATGCACTGAGCTGG 833  
QY 761 CCCATCTGTATCTGTGTCTATCTATCTGATGCTCATGCTGATGCTGAGT 820  
Db 834 CCGCGGTGGAGCTCAACAGCTCGTGGTGTATCTGCTGCTCTACTTCAAGCT 893  
QY 821 GCGCGAGGATCAAGATTCGGCAAGCGGCTACTTCTCGGCTCTTCCATACGT 880  
Db 894 TGGAAAGGCTGAAGACCTCAGGAAAGGTGATGATCAACAGCCATGCTCAT 953  
QY 881 GTCATGATCACTTATATCATCACACATCATCTGCGGCTGCTACTGAGGAT 940  
Db 954 GTCTCTACCTGCTGCTCTGCTGGGCTACCTCTCTGAGCCATGACCGCAT 1013  
QY 941 TCTTGTGTCAGCCTCAATGGGCAACCTCTTGAAGCTGATGATGATGATG 1000  
Db 1014 GCATACCTGAGCTTACCTTCACTGAGGCTCTGAGGAGGCTGTTGGATG 1073  
QY 1001 AGCAGATGTTCTCTCTCTGACAGTGTGACCGGACCATCATGATGTTCT 1060  
Db 1074 ACCGAGGTGTCTCTCTGCTGGGCTGGGCTGGGCTGGTGTGATGCTCT 1133  
QY 1061 AACGTTTTCAGCATATATATCTACAGGATGTTGATGTTGATGATGATG 1120  
Db 1134 AACAAATTCACCAACATCTCTACAGGAGCGATGTCACACCTCCATCACT 1193  
QY 1121 ACAAGTTTCTGCTGTGGGTGACGATCTTCTGATCTTCTGATCCTGGA 1180  
Db 1194 ACAGGCTTCTCTCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 1253  
QY 1181 AACTCAGAGGTGGAGATGATGCTGCTGGGCTGACGATCTGCTTCAATTC 1240  
Db 1254 AGTGTGCTCATCTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1310  
QY 1241 CCTGATGCTATGCTCAAAACATTCACACCTCAGCTATCTGCTGCTGCT 1300  
Db 1311 CCGGAGGACATGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1370  
QY 1301 ATGTGCTGTGGGTATCGGCTATCGCTGCTGCTGCTGCTGCTGCTGCT 1360  
Db 1371 CTGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1430  
QY 1361 ATGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1416  
Db 1431 ATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1490  
QY 1417 CCGT---TTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1474  
Db 1491 GCGACCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1550  
QY 1475 GATGATCACTAGGCTGGAACATCTTCTGCTGCTGCTGCTGCTGCTG 1534  
Db 1551 CTGGAGCAATTTGACCGCGACGCTCATCTCTTGGAGTGTCTCATG 1610  
QY 1535 GTCTTCTGATTTATGATGATGATGATGATGATGATGATGATGATGAT 1594

Db 1611 GTGGCTGTATGATGATGATGATGATGATGATGATGATGATGATG 1670  
QY 1595 AAGAGTGTGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1654  
Db 1671 CCGGCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1730  
QY 1655 GTGTTCTTTCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1710  
Db 1731 GTGTGTGTGTGATGATGATGATGATGATGATGATGATGATGAT 1786  
  
RESULT 2  
US-09-795-232-1  
; Sequence 1, Application US/09795232  
; Patent No. US20010012627A1  
; GENERAL INFORMATION:  
; APPLICANT: Anthony M. Brown  
; APPLICANT: Conrad Gerald Chapman  
; APPLICANT: Israel Simon Glover  
; APPLICANT: Joanne Rachel Evans  
; APPLICANT: William Cairns  
; APPLICANT: Hugh Jonathan Herdon  
; TITLE OF INVENTION: NOVEL COMPOUNDS  
; FILE REFERENCE: GP-30176-D1  
; CURRENT APPLICATION NUMBER: US/09/795,232  
; CURRENT FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 09/182,728  
; PRIOR FILING DATE: 1998-10-29  
; PRIOR APPLICATION NUMBER: 9818890.7  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 2863  
; TYPE: DNA  
; ORGANISM: HOMO SAPIENS  
US-09-795-232-1  
  
Query Match 6.7%; Score 191.6; DB 10; Length 2863;  
Best Local Similarity 52.0%; Pred. No. 6e-31;  
Matches 484; Conservative 0; Mismatches 434; Indels 12; Gaps 2;  
  
QY 746 GGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 805  
Db 1414 GAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1473  
QY 806 GTGTGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 865  
Db 1474 GTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1533  
QY 866 CTCTTCCATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 925  
Db 1534 AGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1593  
QY 926 ACTGAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 985  
Db 1594 GGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1653  
QY 986 TGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1045  
Db 1654 TGGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1713  
QY 1046 ATGTTCTCTCTTACAGGCTTTCAGACATATATATATATATATATAT 1105  
Db 1714 ACTCTCTCTTTCACAAATTTTCACAACTGCTGCTGCTGCTGCTG 1773  
QY 1106 ACTTTGACACCTTTTCAAGTCTTGTGCTGCTGCTGCTGCTGCTG 1165  
Db 1774 TGCACCAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1833  
QY 1166 CTGCGTGAAGATCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1225  
Db 1834 ATGGCCAAATGAACGCAAAATGATGATGATGATGATGATGATG 1890

QY	423	AGTGTGCTCTGGGACATTCAGTTCAAGAAACTCTGTTAAAGTTTGGTCA---ATTTCAC	479
Db	284	AACTGGCCCTGGGCGCAGTACACACCCCTCCGGCTGCTCATTCTCTGGTAAAGGATGTGCC	343
QY	480	CGGCGATGAAGGTCCTGGATACGCTCAAGCTGGCGGGTGGGTTTACATCTCTGATCTACT	539
Db	344	CGGCGCTTAAAGGTCTGGCTATGCTCATCTGCATGATGACATCTACTATGGGCAATGTACT	403
QY	540	ACGTGGTATCTGTTGGTCTCTGTCTGTATTTACTTACGTATGAGCTTCCAGGCCACTCTTC	599
Db	404	ACAACACGATCATGGATGGGCGGTATTTACGTGATGCTTTCTCTCGCGCTATTAACCT	463
QY	600	CATGGGCTATTTGTACGCTGATGGGAACTGGTACCTCTCAGATTCACCACTTGTCTG	659
Db	464	CTGTCTCTCCATGACCACTGCAGTCGCAACAAGTGGAAACACCCGCTGTGCAGCGCGGTCA	523
QY	660	CATCAGTACAAACATTCACCAATGGTTACACAGTGCATCTCACTTTTGTGAACAAG	719
Db	524	CCTCACCCTCAGACTATCTTAACTCTTCTACACCGCGCAAGAGATTTCTTCAGACTATG	583
QY	720	TTCTTCCAACAAAGCGATGGAAATTGAAGAGGTCTCGGTGCCCCCATC---TGGTACTTGG	776
Db	584	TATTTGACACAGCACAAAGTCTAAAGCGCTGATGTACATGAGGCGCATCAAGCGTGG	643
QY	777	TGTTGTCTATTCATGCCATGCGTCATGCTGTTCGGATGGGTGTGGCCCGAGAGATCAAGA	836
Db	644	CTCTGTGTGTGTGGGGTCTTTTGCCTCAGTCACTTCTCTTGTGAAAAGAGATCAAGA	703
QY	837	GTTCGCGAAAGCGGCTACTTCCCTCGGGCTCTTCCCTACCTGATGTATCACTTAT	896
Db	704	GTGCTGGCAAGGTGTGTGGGTGACAGCTCTTGCCCGCTACGTGTGTGCTGTGATTTGC	763
QY	897	TCATCACCACATCATCTCCGCGGTCTACTGAGAGCGATCTGTCTTCTGTCAAGCCTC	956
Db	764	TGGCAGAGAGCGCTACGCTTCACAGAGGAGGAGGATACGCTACTACTTACCCGAG	823
QY	957	AATGGGCAAACTCCTTAGAGCTCGGTGATGTGTACTACGAGTACAGCAAGTGTCTTCT	1016
Db	824	AGTGGACAAATTTGGAAAACTTAAGATGTGATGAGCGCGCATCCGAGATTTTCTTCTCT	883
QY	1017	CTCTGACAGTGTGACCGCGGACCGATCATCATGTTCTCCTCTTAAACAGGTTTCAGACATA	1076
Db	884	CGCTGGGTCCGGGTGTGGGAACCCACTACGGGCGCTCTCAGCTACAAAGTTCACAACA	943
QY	1077	ATATTCACAGGAGATCTTGATTTGTTACGACTTTGGACACTTTCACAGTTTCTGTCTG	1136
Db	944	ACTGCTACAGGAGACCGCTCATCACTTCTTATCAACAGTCTTGACCAAGTCTTCTGCTG	1003
QY	1137	GGTGCACATCTTGGGTATCTCTGTATACCTCGGCTAGCAACTCAACTCAGAGTGGGAG	1196
Db	1004	GTTCCTGATTTTCTCGGTTTGGGTTACATAGGGGCAAGTTTCACAAAG---AAGAGCATCG	1060
QY	1197	ATGTGGTGGGTGCGGCGGTACACAGTCTTGGCTTCATTTCAATTCACCTGATGCCAATTTGCA	1256
Db	1061	AGGAGGTGTGGCTCGAAGGCCCTGACAGTGTGTTCATCTGTATACCCGAGGCATTCGCA	1120
QY	1257	AAACATTTCAAACCTCAGCTATTTTCGGTGTCTTCTCTGATGATGTGCGTGGGTAT	1316
Db	1121	CCATGACGCGGCTCGGTGTGTGGGCAATCTTCTTCTCATGCTTATTAACCTGGGAGAC	1180
QY	1317	TGCGGTATCCGTGGCTGTCTATTCGACTTTCAACACATTGGCGATGAGCGGTTCCAC	1376
Db	1181	TTGACAGTACTTTTGGAGGTCTTGAAGGAGTACACACGCGCTTTTGGCGAATATATCTC	1240
QY	1377	GTGT-----ACCAACCGCTACATATGACAGAGTACCTGTCTTGGGGTTCTCTGCTT	1429
Db	1241	GAGTGTTAGGACAGACATCGCAAGTATTTGTGGCTGTACGTCTTCTGTATCTATATTT	1300
QY	1430	GGACTTGT--TTACTGACACACCGGATGGACAATATATCTTGAGCTTGTAGATCTACAG	1487
Db	1301	GCGCTGTGCCACACACACATACGAGTGTGTATACCTGTGAGACTACTCAATGTGTATG	1360
QY	1488	GTGGAACATTCCTGTGTCTTTTGTGGCCCAATTTCTGAATCTGCAGGCGGTCTTGATTT	1547

Db 1361 GCCCTGATGGCGATCTATTCGTGATTCTGTGAGCGCTGCGGCTGTGCTGGGTGT 1420  
QY 1548 ATGATTTGAGANCTGTGCTAGACATTAGTTGATTTGGTAAAGAACTGTGCTT 1607  
Db 1421 ATGGCGCGACCGGTTCTTGAAGATGTAGAGCAATGCGGACACCCCTGGATGT 1480  
QY 1608 ACTGGGCTCTGCTGGGCGGTATCATCTCTCTATATAGACACTGTGCTTCTACG 1667  
Db 1481 TCTGGAGGACCTTGTGCTTATCATCATGTCCTGATTTCTGTGCTGTGTCT 1540  
QY 1668 CTCTTCTGCTCTTACACACCTGTGTGGAAGAAC 1705  
Db 1541 CCGTCTGCGACACGAGAGATGCTGCGGGAATAC 1578

RESULT 4  
US-10-198-846-11010  
; Sequence 11010, Application US/10198846  
; Publication No. US2003009974A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Steilman, Kathleen  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; FILE REFERENCE: MRI-049  
; CURRENT APPLICATION NUMBER: US/10/198,846  
; CURRENT FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/306,220  
; PRIOR FILING DATE: 2001-07-18  
; NUMBER OF SEQ ID NOS: 14084  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11010  
; LENGTH: 4305  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..2, 4278, 4279, 4280, 4281, 4282, 4283, 4284, 4285, 4286,  
; LOCATION: 4287, 4288, 4289, 4290, 4291, 4292, 4293, 4294, 4295, 4296,  
; LOCATION: 4297, 4298, 4299, 4300, 4301, 4302, 4303, 4304, 4305  
; OTHER INFORMATION: n = A,T,C or G  
US-10-198-846-11010

Query Match 6.58; Score 187.2; DB 15; Length 4305;  
Best Local Similarity 47.78; Pred. No. 6.5e-30;  
Matches 679; Conservative 0; Mismatches 728; Indels 15; Gaps 4;

QY 233 GAGCTGAAACCCGACATGATGATGTCACAAACATGATGATGATGCTGCTG 292  
Db 522 GACCTTCATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 561  
QY 293 ATGCTACATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 352  
Db 582 ATGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 641  
QY 353 GAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 412  
Db 642 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 701  
QY 413 TACTACTAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 469  
Db 702 TTTTACATGAGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 761  
QY 470 TCAATTTACCGGCAATGAAGGTACTGATACGCTCAAGCTGCGGCTGCTACATC 529  
Db 762 AAATATGCGCCGATTTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 821  
QY 530 CTGTCTTACTAGTGTGAGATGCTGTCTGTCTGTATCTTACTAGTATGAGCTCCAG 589

Db 822 GCTTCTACTACAAACATCATGAGGCTGGGCGCTATCTACCTCATCTCTCTCAGC 881  
QY 590 GCCACTTCCATGAGGCTATTTGACGCCGAGTGGGAGAGACGCGCATACCTCAGATCCA 649  
Db 882 GACAGTGGCTGGAGACACCTGACCAAGACTCTGGAACCTGGAACCTGACCAATTAC 941  
QY 650 ACATTTGCTCATACGACCAACATCAGCAATGATGACGAGCTGCACTCACTTCT 709  
Db 942 TTTCTCGAGAGCAACATCAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1001  
QY 710 TTGAGAACAGTTTCCA--ACAAGGATGAAATGGAAGAGGCTGCTGCTGCTGCTG 766  
Db 1002 ACCGCGACCTCTGAGATCCACCGGTAAAGGAGGCTGAGGAGGAGGAGGAGGAG 1061  
QY 767 TGTACTTGTGTGTGTCTATTCATTCGATGCTCATGCTGTGCTGAGTGTGCTGCTG 826  
Db 1062 TGGAGCTGGGCTCTGACATCATGATGCTGATCTGATCTGATCTGATCTGATCTG 1121  
QY 827 GAGTCAAGAGTTCGCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 886  
Db 1122 GGGTCAAGACCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1181  
QY 887 ATCACTTATTCATCACCACATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 946  
Db 1182 TCTGTCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1241  
QY 947 GTACGCTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1006  
Db 1242 TTGAAACCAATTTGCGAAGAACTCTGAGACAGGAGGAGGAGGAGGAGGAGGAGGAG 1301  
QY 1007 GTGTCTCTCTGACAGTGTGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1066  
Db 1302 ATCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1361  
QY 1067 TTGAGCATATATATACAGAGGATGCTGATGATGATGATGATGATGATGATGATG 1126  
Db 1362 TTCAACAAACATGCTACCAAGATGCTGATGATGATGATGATGATGATGATGATG 1421  
QY 1127 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1186  
Db 1422 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1241  
QY 1187 GAGTGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1246  
Db 1482 GATGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1541  
QY 1247 GCCATGCCAAACATTCACACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1306  
Db 1542 GCGATAGCCAAACATGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1601  
QY 1307 GTGCTGGGATGCGGCTCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1366  
Db 1602 AGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1661  
QY 1367 GCGTTCACAGTGT-----ACCAACGCTTACATGTCAGGATGACCTGTTCTT--GC 1417  
Db 1662 GAGTTCACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1721  
QY 1418 GGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1477  
Db 1722 TGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1781  
QY 1478 GATCACTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1537  
Db 1782 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1841  
QY 1538 TTTCTGATTTATGATTTGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1597  
Db 1842 TCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1901  
QY 1598 ACTGAGTCTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1639  
Db 1902 CCGGGGAGGTTCTGAGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1943







US-09-954-456-560

Query Match 5.6%; Score 160.6; DB 11; Length 2283;  
Best Local Similarity 46.3%; Pred. No. 2.4e-24;  
Matches 722; Conservative 0; Mismatches 789; Indels 48; Gaps 4;

QY 260 TGTGCAACAACATGTAATCTTGTATGCTGCAATCGCTACATCGCGTGGTGGTAAAC 319  
DB 181 TGGACAGCCAGATGAGACTCATGCGTGGGCTTCCGCGTGGGCTGGTAAAC 240  
QY 320 GTGGGGGCTCCCTTTCATCGCTACAGAAATGAGAGAGTCTTCTCGTGGCAATAC 379  
DB 241 GTGGGGGCTCCCTTTCATCGCTACAGAAATGAGAGAGTCTTCTCGTGGCAATAC 300  
QY 380 GTACATGCTTTTACTGTCGGCAAGCGCTGTACATCTAGAGTGTGCTGGACAA 439  
DB 301 GTCTGATTTGCCCTGGTGGAGAAATCCCATTTTCTTCTGGAATCTCATCTGGCCAG 360  
QY 440 TTCAATTCAGAAACCTGTTAAAGTTGTCAATTCACCGGCCATGAAGTACTGGA 499  
DB 361 TTCAATTCAGAAACCTGTTAAAGTTGTCAATTCACCGGCCATGAAGTACTGGA 420  
QY 500 TACCTCAAGTCCCGGCTGGGTTACATCTGCTTACTACGCTGATCTGTGCTC 559  
DB 421 TATACCTCCAGATGATGTTCTTCTACGCAACACTTACTATCATGCTGCTGGCTG 480  
QY 560 TGTGTAATTTACTAGTATGAGCTTCCAGGCCACTCTTCCATGGGCTATTGTGAGCT 619  
DB 481 GGTCTTCTATTTACCTGGTCAATGCTTCTTACTACACTTTGCCATGGGCTAGCTGGCCAC 540  
QY 620 GAGTGGGAAATCGCTGACCTCAGATCCAACTTGTGCTATCAGTCAAGCAATACACC 679  
DB 541 ACCTGGAACACTCTGATGATGAGATCTTTGACATGAAGACTGACCAATACAGC 600  
QY 680 AATGTAACCAAGTGCACAC-----TCTACTTTTG 712  
DB 601 TTGGCAACCTCACATGTCAGCACTGTCAGCCGGGCTCCCTGTATGAGTGTGG 660  
QY 713 AGAACATTTCTCAACAAGGATGATGAAGAGTCTCGGCGCCCATCTGTGTAC 772  
DB 661 GAGAACAAAGCTTGGAGGCTCTCCACAGGCTGGAGTTCCAGAGGCCCTCACTGAGGAG 720  
QY 773 TTGGTGTGCTTATTCATGATGATGCTCATGCTGATGCTGAGTGTGCGCCGAGAGT 832  
DB 721 GTGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
QY 833 AAGAGTCCGCAAGGCGGCTACTCTCTGCGGCTCCCTCCATAGTGTGATATCACT 892  
DB 781 AATTCACGGGAAAGATGCTGTACTCTCACTCTACATTTCCCTACGTGCTGTGCTG 840  
QY 893 TTATTCATCACCAATCATCTGCGGCTGCTACTAGGCAATCTTCTTCTGCTACAG 952  
DB 841 CTGCTGGTGGTGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
QY 953 CCTCAATGGGCAACTCTTGAAGTCTGATGATGATGATGATGATGATGATGATGATG 1012  
DB 901 CCTCAATGGGCAACTCTTGAAGTCTGATGATGATGATGATGATGATGATGATGATG 960  
QY 1013 TTGCTCTGACAGTGTGACCGGACCGATCATGCTCTCTCTTAACAGGTTTCA 1072  
DB 961 TTGCTCTGACAGTGTGACCGGACCGATCATGCTCTCTCTTAACAGGTTTCA 1020  
QY 1073 CATATATCTACAGGATGCTTGTGATGATGATGATGATGATGATGATGATGATGATG 1132  
DB 1021 AACAACATCTACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
QY 1133 TCTGAGTGCAGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1192  
DB 1081 GCTGAGTGCAGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140  
QY 1193 GAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1252  
DB 1141 TCCAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1197

QY 1253 GCCAAACATTCACACCTACGATATCTCGGAGCTGTTCTTCCATGATGATGAGGCTG 1312  
DB 1198 ACACGATGCTGTGGGCCCTACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1257  
QY 1313 GGTATGCGCTACCGTGGCTCTGCTATCGATTCGATTCACATTCGAGATGAGGAGCGCTTC 1372  
DB 1258 GGTGAGACAGCATTTGTAGTGTGAGGAGCTTCATCACATGCGGCTCTGATGCTGCTC 1317  
QY 1373 CCA-----CGTGTACCCACCGTCTACATGATGATGATGATGATGATGATGATGATG 1417  
DB 1318 CCGGCTCTCTACTCTTCCGTTTTCAAAGGAGATCTCGGCGCTCTGTTGCGCTC 1377  
QY 1418 GATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1477  
DB 1378 TGTCTTGTATCATCTCTCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 1437  
QY 1478 GATCATACGAGTGAACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1534  
DB 1438 GATCATACGAGTGAACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1497  
QY 1535 GTGTTCTGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1594  
DB 1498 GTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1557  
QY 1595 AAGACTGGTGTACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1654  
DB 1558 GCACCTTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1617  
QY 1655 GTGTTCTTCTACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1714  
DB 1618 ATCTCATCTTCAACGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1677  
QY 1715 CCGACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1773  
DB 1678 CCGTGTGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1736

RESULT 8  
US-09-954-456-1630  
Sequence 1630, Application US/09954456  
Patent No. US20020115057A1  
GENERAL INFORMATION:  
APPLICANT: Young, Paul  
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C  
TITLE OF INVENTION: Sets  
FILE REFERENCE: 689290-76  
CURRENT APPLICATION NUMBER: US/09/954,456  
CURRENT FILING DATE: 2001-09-18  
PRIOR APPLICATION NUMBER: US/60/233,617  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: US/60/234,052  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: US/60/234,923  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: US/60/235,134  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: US/60/235,637  
PRIOR FILING DATE: 2000-09-26  
PRIOR APPLICATION NUMBER: US/60/235,638  
PRIOR FILING DATE: 2000-09-26  
PRIOR APPLICATION NUMBER: US/60/235,711  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,720  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,840  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,863  
PRIOR FILING DATE: 2000-09-27  
NUMBER OF SEQ ID NOS: 2276  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 1630  
LENGTH: 2283

TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc\_feature  
 OTHER INFORMATION: n=a,t,g or c  
 US-09-954-456-1630

Query Match 5.68; Score 160.6; DB 11; Length 2283;  
 Best Local Similarity 46.3%; Pred. No. 2,4e-24;  
 Matches 722; Conservative 0; Mismatches 789; Indels 48; Gaps 4;

QY 260 TGGTCCACAAACATGTAATCTGATGCTCTGATCGCTACATCCGCTGGTTGGGTAC 319  
 DB 181 TGGACACGCGAGATGACTTATCATCTGCGTGGCGCTGCGGCGCTGGCTGATAC 240  
 QY 320 GTGTGGCGGTCCCTTCATCGGCTACCAAGATGAGAGAGTGGCTTCCTGATCCATAC 379  
 DB 241 GTGTGGCGGTCCCTTCATCGGCTACCAAGATGAGAGAGTGGCTTCCTGATCCATAC 300  
 QY 380 GTCATCGTCTTTTACTTGTGGCAAGCCTGTGTACTACTTGAAGTGTCTCGACAA 439  
 DB 301 GTCTGATTCCTCGTGGAGGATCCCATTTTCTTCTGGAAATCTCATGCGGACAG 360  
 QY 440 TTCAAGTTCAGAAACCTGTAAAGTTGGCAATTCACCGGCGCATGAAGGTACTGA 499  
 DB 361 TTCAAGTTCAGAAACCTGTAAAGTTGGCAATTCACCGGCGCATGAAGGTACTGA 420  
 QY 500 TACGCTCAAGCTCCCGGCTCGGTACATCTGTCTTACTACTGATGATCTGTGCTC 559  
 DB 421 TATGCTCCATGGATGTCTTCTACTGCAACACTTACTACATCATGATGATGCTGCTG 480  
 QY 560 TGTCTGATTTACTTACTGATGATGATGATGATGATGATGATGATGATGATGATG 619  
 DB 481 GGTCTTATTTACTTACTGATGATGATGATGATGATGATGATGATGATGATGATG 540  
 QY 620 GAGTGGAGAACTGCGTACCTGATGATGATGATGATGATGATGATGATGATGATG 679  
 DB 541 ACCTGGAGAACTGCGTACCTGATGATGATGATGATGATGATGATGATGATGATG 600  
 QY 680 AATGTACACAGAGTGTCTAAC-----TCTACTTTTG 712  
 DB 601 TTGGCCAACTCATATGATGACAGCTTGTGACCGGCGGTCCCTGTCTCATGAGTTCTG 660  
 QY 713 AGAAGATTTCTCCAAACAAAGCGATGATGATGATGATGATGATGATGATGATGATG 772  
 DB 661 GAGAACAACTCTTGAAGCTCTCACAGGCTGAGAGTTTCCAGAGCCCTCAACTGAGAG 720  
 QY 773 TTGCTGTGTCTTATTCATGCGATGATGATGATGATGATGATGATGATGATGATG 832  
 DB 721 GTTACCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
 QY 833 AAGAGTTCCGGAAGGCGCTTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 892  
 DB 781 AATTCACAGGGAAGATCGTACTTACTTACTTACTTACTTACTTACTTACTTACTT 840  
 QY 893 TTATTCATCACCAATCATCTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 952  
 DB 841 CTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
 QY 953 CCTCAATGGGGAAGACTCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1012  
 DB 901 CCTGACTGTGCAAAAGCTGGGCTCCCTCAAGGTGTGATGATGATGATGATGATGATG 960  
 QY 1013 TTCTCTCTGACAGTGTGACCGGACGATCATCTTCTCTCTCTCTCTCTCTCTCTCT 1072  
 DB 961 TTCTCTTATGCGATGCGGCTGGGCGGCTGACAGCCCTGGGACGATACATCTCTTCAAC 1020  
 QY 1073 CATTAATATCTACAGGATGCTGATTTTATGACATTTTGGACACTTTTACAGTTTCTTG 1132  
 DB 1021 AACAACTGTCTACAGGATGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
 QY 1133 TCTGGGTGACAGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1192

DB 1081 GCTGGCTTTGTGCTTCTCCATCTCGGCTTCATGCGCACAGAGAGGTGCTCATATC 1140  
 QY 1193 GGAATGTGGTGGTGGTGGTGGGACAGCTTCTGCTTCAATTCATACCTGATGACAT 1252  
 DB 1141 TCCAGGTGGAGATAGAGGCTG---GCTTACCTTCACTGCTACACAGGCGGCTC 1197  
 QY 1253 GCCAAACATTCACACCTCAGCTATTCGCTGCTGCTTCTTCCATGATGATGCTGCTG 1312  
 DB 1198 ACATGATGCTGTGGGCGCCACCTGTGGGCTGCTGCTGCTTCTTCTTCACTGCTGCTC 1257  
 QY 1313 GGTATGCGCTATCCGCTGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 1372  
 DB 1258 GGTCTGACACCGAGTTGTAGTGTGAGGCGCTTCACTACCTGCGCTCTGATCTCTC 1317  
 QY 1373 CCA-----CGTATCCACCGTGTACATGTCAGGAGTACGATCTGCTTCTG 1417  
 DB 1318 CCGGCTCTCTACTTCTCCGCTTTTCAAGGAGATCTCCGCTGCGCTCTGCTGCTC 1377  
 QY 1418 GGTTCCTGCTTGGACTTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1477  
 DB 1378 TGTCTTCTATGCTATCTCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1437  
 QY 1478 GATCTACAGGTGACATTCCTGCTTCTTCTGCT---GCCATTTCTGAACTGCGAGC 1534  
 DB 1438 GACTACTACAGCTAGTGTGCTACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1497  
 QY 1535 GTTCTGATGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1594  
 DB 1498 GTGGCTGGGTATGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1557  
 QY 1595 AAGACTGTGCTTACTGCGCTCTCTGCTGCGGCGTAACTCTCTCTATTAATGACACT 1654  
 DB 1558 CGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1617  
 QY 1655 GTTCTTCTACAGCTCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1714  
 DB 1618 ATCTTATCTTCAACAGTTGTGTACTACGACCGCTGCTGCTTCAACAACACCTTAC 1677  
 QY 1715 CCGACTGCTGCTTATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1773  
 DB 1678 CCGTGTGGGCTGAGGCGCATGGGCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTG 1736

RESULT 9  
 US-09-919-039-377  
 : Sequence 377, Application US/09919039  
 : Publication No. US20030108871A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Kaser, Matthew R.  
 : TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES  
 : FILE REFERENCE: PA-0035 US  
 : CURRENT APPLICATION NUMBER: US/09/919,039  
 : CURRENT FILING DATE: 2002-09-09  
 : PRIOR APPLICATION NUMBER: 60/222,113  
 : PRIOR FILING DATE: 2000-07-28  
 : NUMBER OF SEQ ID NOS: 401  
 : SOFTWARE: Perl Program  
 : SEQ ID NO 377  
 : LENGTH: 2919  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : FEATURE:  
 : NAME/KEY: misc\_feature  
 : OTHER INFORMATION: Incyte ID No. US20030108871A1 5834958CB1  
 US-09-919-039-377

Query Match 5.48; Score 154.8; DB 12; Length 2919;  
 Best Local Similarity 46.5%; Pred. No. 4.9e-23;  
 Matches 738; Conservative 0; Mismatches 812; Indels 36; Gaps 6;  
 QY 221 GACACGACTTAAAGCTGACCGCCAGACGATGATGATGATGATGATGATGATGATG 280  
 DB 136 GACACGAGACGAGACCGAGGTGAAGATGCGGCGCAATGACCAACAGATGAGATT 195

```

OY 281 TTGATGTCATGATCGCTACATCGGTGTTGGGTAAACGTTGGGGTTCCTTCATC 340
  || || || || || || || || || || || || || || || || || || || || ||
Db 196 GTGCTGTCATGAGCGCGGAGATCATTTGGCTGGGCAATGCTGGAGTTCCCTATC 255
OY 341 GCGTACCAGATGAGAGAGGTGCTTCTGTTGTCATACGTCATGCTTCTTCTTCTGTC 400
  || || || || || || || || || || || || || || || || || || || || ||
Db 256 TGTACAAAACGAGAGTGGAGCTTCTCATGCCCTTCTCATCTTCTTCTTCTGTC 315
OY 401 GCGAGCTGTGTCATGTTAGAGTGTGCTCGGCAATTCAGTTCAAGAACTGTT 460
  || || || || || || || || || || || || || || || || || || || || ||
Db 316 GGGATCCCGGTCTTCTGAGAGTGGCGTTGGGCAATACACCAAGGAGAGTGC 375
OY 461 AAGTTT---GTCATTTACCGGCGCATGAANAGTACATGATAGCTCAAGTCCCGGC 517
  || || || || || || || || || || || || || || || || || || || || ||
Db 376 ACAGCTGAGAGAAATGCTCCCTCTTCCAGGCAATTTGGTCTGAGTCTGTGTCATC 435
OY 518 TGGGTTACATCTGCTTCTTCTTCTGATGATGATGATGATGATGATGATGATGATG 577
  || || || || || || || || || || || || || || || || || || || || ||
Db 436 GAGTCATTTTGAATGTCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 495
OY 578 ATGAGCTTCCAGGCGCTTCCATGAGGCTATTTTGTACGCTGAGTGGAGAACTGCTGA 637
  || || || || || || || || || || || || || || || || || || || || ||
Db 496 AGCTCTTCACTCTGAGCTGCTGGAGCACTGCAACAACTTTTGGAAACAGAGCAT 555
OY 638 CCTCAGAT-----CCAGACTTGTGTCATGATCAACAAACATCACCAATGATGATG 688
  || || || || || || || || || || || || || || || || || || || || ||
Db 556 TGCAGGAGACTTCTGAAACACTCAGAGCGGCGGCACTGACCCCTTTGAGAAATTTTACC 615
OY 689 AGCAGTCTCACTACTTTTGGAGAAAGTTCTCCAAACAAAGATGAAATGAAAGA 748
  || || || || || || || || || || || || || || || || || || || || ||
Db 616 TCACCTGTCAATGAAATTTGGGAGAGAGAGTTCTGGGCAATCCTCGGGCATCA--T 672
OY 749 GGTCTGGTGGCCCGCATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 808
  || || || || || || || || || || || || || || || || || || || || ||
Db 673 GACTTGGCTCCCTGCGCTGGAGACTGGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 732
OY 809 TTGGAATGCTGCGCCGAGAGTCAAGATTTCCGGAAGCGGCTACTTCTTCTGCTGCTG 868
  || || || || || || || || || || || || || || || || || || || || ||
Db 733 TATTTTCGATCTGAGAGGGGTTCAATCCACAGGCAAGGTTTATTTTACAGCCACG 792
OY 869 TTCCATAGCTTGTGATGATCACTTTATATATCAACCAATCAATCTCCCGGTGCTACT 928
  || || || || || || || || || || || || || || || || || || || || ||
Db 793 TTTCCTGATCTGATGTTGTTGATTTGCTATGATCAGAGTGCATCCCTCCGAGGCTTAC 852
OY 929 GAGGGAATCCTGTTCTGTCACGCCCTCAATGGCGGAACTCCTTGAAGTCTGGTATG 988
  || || || || || || || || || || || || || || || || || || || || ||
Db 853 CAGGGCAATCACTACTACTGGAAGCAGATTTGTTCCGCTCAAGGACCTCAGGTGTGG 912
OY 989 TACTCAGACGTCAAGCAAGTGTCTTCTCTGACAGTGTGACACCGATCATCATG 1048
  || || || || || || || || || || || || || || || || || || || || ||
Db 913 ATGATGCGGGGACACCGATCTTCTCTCTTGGCATCTGCGAGGGGTGCTACAGACC 972
OY 1049 TTCTCTCTTACAAAGGTTTACAGATATATATCTACAGGAGTGTGGATTTTACGACT 1108
  || || || || || || || || || || || || || || || || || || || || ||
Db 973 CTGGGCGCTACACAAAGTATACAAACACTGTCTACAGGACTCATGCGCTTGTCTTC 1032
OY 1109 TTGGACACCTTTTCAAGTTTCTGTGGTGGGACAGATCTTGGTATCTTGTACCTC 1168
  || || || || || || || || || || || || || || || || || || || || ||
Db 1033 CTGACAGCTGCCACACACTTGTGGTGGGTTGTTGTTCTTCTTCTTCTTCTTCTTCTTCT 1089
OY 1169 GCGTACAGACTCAAGAGTGGAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1228
  || || || || || || || || || || || || || || || || || || || || ||
Db 1090 ATGTCACAAAGAGAGGAGTGGCCATTTTGAAGTGGCGAGTGTGCTGCTGCTGCTGCTG 1149
OY 1229 TTGATTTTCAATCTGATGCTATTTGCCAAAACATTTCAACCTCACTATTTCTGCTGCTG 1288
  || || || || || || || || || || || || || || || || || || || || ||
Db 1150 TTGATTCCTTCTTCCCAAGGCTGTGACTATATATGCTTATTCAGCTGTGCTGCTGCTG 1209
OY 1289 TTCTTCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1348
  || || || || || || || || || || || || || || || || || || || || ||
Db 1210 TTCTTTTATCATGCTCATATTTCTTATGAGGCTGACAGCCAGTTTGTCTGTGTGAGTGCCTG 1269
```

```

OY 1349 AACCATTTGGGATGAGAGCGGCTTCCAGGTGATCCAC-----GTCATC 1393
  || || || || || || || || || || || || || || || || || || || || ||
Db 1270 GTGACAGCTTCAATAGCATATGTTCCAGGAGCTCCGGAAGAGCGGCGGAGCTC 1329
OY 1394 ATGTCAGAGATGAGCTGTTCTTGGGTTTCTGCTTGGAGCTTGTATCTGACACCGAGT 1453
  || || || || || || || || || || || || || || || || || || || || ||
Db 1330 CTGATCTCAACCATCGCGGCTATGCTTACTACTGATAGGCTTTTCTGCTGCTGCTGCTG 1389
OY 1454 GGCATATATATTTTGGACTTGTATGATCACTAGGTTGGAACATTTCT--TGTCTTTTC 1510
  || || || || || || || || || || || || || || || || || || || || ||
Db 1390 GGGATGTATACATCTTCCAGCTGTTGATGATGATGATGATGATGATGATGATGATGATG 1449
OY 1511 TGGCCATTTCTGAACTCGAGCGGCTGTTCTGATTTATGATTTGAGAACTGCTGCTA 1570
  || || || || || || || || || || || || || || || || || || || || ||
Db 1450 CTGATATTTTGAAGTGTGCTGATAGCTGGGTTGATGAGGCGGACCGCTTCTATGAC 1509
OY 1571 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1630
  || || || || || || || || || || || || || || || || || || || || ||
Db 1510 AACATTTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1569
OY 1631 ATCACTCCTGCTATATATGAGACTGTGTTCTTACGCTCTTCTGCTCTTACACCTG 1690
  || || || || || || || || || || || || || || || || || || || || ||
Db 1570 CTGACCCCTGAGCTTGGCTGCGCCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1629
OY 1691 GTGTTGGAAGCACTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1750
  || || || || || || || || || || || || || || || || || || || || ||
Db 1630 AAGTACAAACAGCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1689
OY 1751 TTTTGGGATGAGCTTGTGCAAT 1776
  || || || || || || || || || || || || || || || || || || || || ||
Db 1690 CTGCTCCATGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1715

RESULT 10
US-09-880-107-3339
: Sequence 3339, Application US/09880107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherf, Uwe
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921-5028-MO
: CURRENT APPLICATION NUMBER: US/09/880,107
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 3339
: LENGTH: 3410
: TYPE: DNA
: ORGANISM: Homo sapiens
: OTHER INFORMATION: Genbank Accession No. US20020142981A1 U27699
US-09-880-107-3339

Query Match 5.4%; Score 154.8; DB 11; Length 3410;
Best Local Similarity 46.5%; Pred. No. 5.3e-23;
Matches 738; Conservative 0; Mismatches 812; Indels 36; Gaps 6;

OY 221 GACACGACTTAGAGCTGAACCGCCAGAACATGATGATGATGATGATGATGATGATGATG 280
  || || || || || || || || || || || || || || || || || || || || ||
Db 662 GACACGAGAAAGAGAGGAGCAGAGTGAAGATGCGGGCCAAATGAGCAACAGATAGAGTTT 721
OY 281 TTGATTCCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 340
  || || || || || || || || || || || || || || || || || || || || ||
Db 722 GTGCTGTACAGTGGCGGAGAGATCATTTGGGCTGGCAATGCTGAGAGTTTCCCTATCTC 781
OY 341 GCGTACCAAGATGAGAGAGTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 400
```

DB 782 TCGTACAAAACGAGGTGGAGGCTTCTCATCCCTCACTTCACTTCTTGTGTGCG 841  
QY 401 GGCAGAGCTGTGACTACTAGAGTGTCTCGGCAATTCAGTTCAAAACTGCTT 460  
DB 842 GGCATCCCGGTGTCTTCTCGAGAGTGGCTTGGCCAAATACACCAAGGAGGTG 901  
QY 461 AAGATT--GCTCAATTTCCAGGCGCATGAAAGTACTGATACGCTCAAGCTGCGGC 517  
DB 902 ACAGCCTGGAGAAATCTCCCTCTTCCAGGGCATGTGCTGGCATCTGTGTATC 961  
QY 518 TGGGTTTACATCTCTGTCTTCTAGAGTGTATCTGTGTCTGTCTGTATTAAGCT 577  
DB 962 GACTCATATTGTAATGTCTACTATCATCATCTTGGCTGGGCTCTCTTCTACCTGTT 1021  
QY 578 ATGAGCTTCCAGGCGCACTTTCATGGGCTATTTGTGAGGCTGAGGAGAACTGGGTA 637  
DB 1022 AGCTCTTCACTTCTGAGCTGCGCTGACACCTGCAACAACCTTTGGAACACAGACAT 1081  
QY 638 CCGTCAGAT-----CCACACCTTGTGATCATCATCAACAACATCAACAAATGTACC 688  
DB 1082 TGCACGAGCTTGTGAAACCACTCAGAGAGCGGACAGATGACCCCATTTGAAATTTTACC 1141  
QY 689 AGCAGTCTCACTACTTCTTTGAGAACGTTCTCCAAACAACGATGGAATGAAGA 748  
DB 1142 TCACCTGTCAATGGAATCTGGGAGAGACGAGTCTGGGCAATCACTGGGCACTCA--T 1198  
QY 749 GGTCTGGGTGCGCCCACTGCTACTTGTGTGATCTATTCATGCGATGGCTCATGCTG 808  
DB 1199 GACTTGGGCTCCCTGGGCTGGAGCTGGCCCTGTGCTCTGCTGCTGGCTGGCTATCTGC 1258  
QY 809 TTGCGAGTGTGCGCCGAGAGTCAAGAGTTCGCGCAAAAGCGGCTACTTCTCGGCTC 868  
DB 1259 TATTTCTGCTGTGAAGGGGCTCAAGTCCACAGGCAAGGTGGTTATTTACAGCGACG 1318  
QY 869 TTCCCATGAGTGTGATGATCACTTTATTCATACACAAATCATCTGCGCGGCTACT 928  
DB 1319 TTTCCGTAAGTGTGCTGCTGATTTGATGATGAGGTGTCACCTTCCCGAGGCTTAC 1378  
QY 929 GAGCGATCGTGTCTTCTGTCAGCGCTCAATGGGCAAACTCTTAGCTGGGTATG 988  
DB 1379 CAGGGCATCATCTACTGTAAGGCAAGATTTTCCGCTTCAAGACCTCTCAGGTG 1438  
QY 989 TACTCAGAGTCAAGAGTGTCTCTCTGACAGTGTGACCGGACGATCATCANG 1048  
DB 1439 ATGAGTGGCGGCGACCCAGATCTTCTCTTCTGATCTGCGAGGGGTCTGACAGCC 1498  
QY 1049 TTCTCTTACAGAGGTTTCAAGATTAATATCTACAGGATGCTTGTGATGTAGACT 1108  
DB 1499 CTGGGCAAGCTACACAAGTATCACAAACAGTCTACAGAGATGATCGCCCTGTGCTTC 1158  
QY 1109 TTGGACACCTTTACAAGTTTCTTGTGGTGCACATCTTGGTATCTGTGTAACCTC 1168  
DB 1559 CTGAACAGTGGCCACAGCTTGTGGGTGGTGTGTCTTCTTCCATCCGGG---CTTC 1515  
QY 1169 GCCTACAGACTCACTCAGAGTGGAGATGTGTGCTGGTGGCGGTACAGTCTTCT 1228  
DB 1616 ATGTCCCAAGAGCAAGGGGTGCCATTTCTGAAGTGGCGAGTCAAGTCTGGCTGGCC 1675  
QY 1229 TTCAATTTCAACCCCTGATGCGATTTGCCAAAACATTCACAACTGATCTTGTGGTCTG 1288  
DB 1676 TTCAATGCGCTTCCCAAGGCTGTGACTATGATGCCCTTATTCAGAGTGTGCTGCTC 1735  
QY 1289 TTCTTCTGATGATGTGCGTGTGCTATCGGCTATCGGTGCTGTGATGACTTTC 1348  
DB 1736 TTCTTATCATGTGCTATTTCTCAGGGCTGGACAGCAAGTGTCTGTGTGGAGTGGCT 1795  
QY 1349 AACCAATTTGGCGATGAGCGCTTCCACAGTGTACCCAC-----GCTCTAC 1393  
DB 1796 GTGACAGCTCCATACAGATGTTCCTCCAGAGCTCCGAAAGAGCGGCGGCGAGCTC 1855  
QY 1394 ATGTACGAGATGACCTGTCTTCTGCTTCTGAGTCTTGTACTGCAACCGGGT 1453

DB 1856 CTCATCTCACCAATCGCCGTGATGTGCTACTGATAGGCTTTTCTGTACCGAGGC 1915  
QY 1454 GGAACAATATTTCTGTAGCTTTGATGACTACTAGGTGGAACATTCCT---TGTGCTTTC 1510  
DB 1916 GGGATGTACTATCTTCCAGCTGTTGTGACTACTATGCTTCCAGTGGCAATTCCTGTCT 1975  
QY 1511 TGGCCATTTCTGAACCTGCGAGCGTGTGTGATTTATGATTTGAAATCTGTGCTTA 1570  
DB 1976 CTGTGATGTTTGAAGTGTGTGATGATGAGTGGGTGTATGGGGGAGACGCTTCTATGAC 2035  
QY 1571 GACATGAGTTCATGTTGGGTAAAGACTGTGTCTTACTAGCGCTCTCTCTGGGCGTA 1630  
DB 2036 AACATTGAGGACATGATTTGGCTACCGGCTAGGCGCTGTGTGAAGATCTCTGCTTTC 2095  
QY 1631 ATCACTGCTCTTAATGAGAGCTGTGTCTTCTAGCCTCTTCTCGGCTCTAACACCTG 1690  
DB 2096 CTGACCCCTGACTTTCCTGCGCACCTTCTCTTCTTCTTGTGACAAATGACACCCCTC 2155  
QY 1691 GTGTTGGAGACAATACATGATACCCGACGTGTGTATGTTTCTGATACTTGATGTA 1750  
DB 2156 AAGTACAAACAGCTATGTGTACCGGCTGGGAGATACATGATGCTGTGCTGTGCT 2215  
QY 1751 TTTTGGGCAATGACGTTTGTCCAT 1776  
DB 2216 CTGTCTCATGAGTGTGTGTGCTCCACT 2241

## RESULT 11

US-09-917-800A-1429  
Sequence 1429, Application US/09917800A

Patent No. US20020119462A1

GENERAL INFORMATION:

APPLICANT: Mendrick, Donna

APPLICANT: Porter, Mark

APPLICANT: Johnson, Kory

APPLICANT: Castle, Arthur

APPLICANT: Elashoff, Michael

TITLE OF INVENTION: Molecular Toxicology Modeling

FILE REFERENCE: 44921-5038-US

CURRENT APPLICATION NUMBER: US/09/917, 800A

CURRENT FILING DATE: 2001-07-31

PRIOR APPLICATION NUMBER: US 60/222,040

PRIOR FILING DATE: 2000-07-31

PRIOR APPLICATION NUMBER: US 60/222,880

PRIOR FILING DATE: 2000-11-02

PRIOR APPLICATION NUMBER: US 60/290,029

PRIOR FILING DATE: 2001-05-11

PRIOR APPLICATION NUMBER: US 60/290,645

PRIOR FILING DATE: 2001-05-15

PRIOR APPLICATION NUMBER: US 60/292,336

PRIOR FILING DATE: 2001-05-22

PRIOR APPLICATION NUMBER: US 60/295,798

PRIOR FILING DATE: 2001-06-06

PRIOR APPLICATION NUMBER: US 60/297,457

PRIOR FILING DATE: 2001-06-13

PRIOR APPLICATION NUMBER: US 60/298,884

PRIOR FILING DATE: 2001-06-19

PRIOR APPLICATION NUMBER: US 60/303,459

PRIOR FILING DATE: 2001-07-09

NUMBER OF SEQ ID NOS: 1740

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1429

LENGTH: 2028

TYPE: DNA

ORGANISM: Rattus norvegicus

FEATURE:

OTHER INFORMATION: Genbank Accession No. US20020119462A1 M95762

US-09-917-800A-1429

Query Match 5.2%; Score 147.6; DB 11; Length 2028;  
Best Local Similarity 46.7%; Pred. No. 1,4e-21;  
Matches 699; Conservative 0; Mismatches 764; Indels 33; Gaps 6;

OY	248	GAAGTATGATGGTCCAAACAATGTAATCTTGATGCTCGATCGCTACATCCGTC	307
Db	216	GAAGGGAGCAATGATGACCACCAAGATGGAGTTGCTGATGAGTGGGGAGAGATCAAT	275
OY	308	GATTGGGTAACGTGTGGCGGTTCCCTTTCATCCGCTACCGAAATGAGAGAGTCTTTC	367
Db	276	GGCTTAGGCAACGCTGGAGGTTCCCATCTGCTGCTACAGAAAGGGGGAGTGGCTTC	335
OY	368	CTGGTGCATACGTCATCTGTTCTTTTACTTGTGGCAAGCCTGTGTACTTACATGTG	427
Db	336	TTTATATCCCTACCACTTCCTTTCATTTACGTGGCAATCCGTCCTTCTCTGGAGACA	395
OY	428	GTCTCGGACATTCAGTTCAAGAAACTCTGTAAAGTT---GGTCAATTCACCGGAC	484
Db	396	GGCGTTGGCCAGTACACCAACGAGGAGCATACAGGCTGGAGAAATCTGTCCATC	455
OY	485	ATGAAGAAGTACTGATACGCTCAAGAGCTGGCGGTGCTACATCTGTCTTACTAGTG	544
Db	456	TTTCAGGGCAACGGCTATGACCTCACAGATGATGTCACCCCTTTCATATGTCTACTACATC	515
OY	545	GTGATCTGTGTCCTGCTGTATTTACTTACTTATGACCTTCACGGGCACTGCTCCATAG	604
Db	516	GTGTCTCCGCGCCTGGGCCCTTTTACTCTTTCAGAGCTTCACACACTGACCTCCCTGG	575
OY	605	GCTATTGTCAGCCTGAGTG-----GAGAACTGCGTACCCCTGAGATCAACACTTGCT	658
Db	576	GGTAGCTGCACGCCAGATGGAAATACAGAAAACGTGTGGAGTTCAGAAAACCAACAT	635
OY	659	GCATCAGTCAACAACATCACCAATGGTATGATCCAGAGCTGTCAACTCTTTTGGAGACA	718
Db	636	TCCCTGATGATGACTTCTGAGAAATGCCACATCCCTGTCAATGAGTTGTGGAGGGCGA	695
OY	719	GTCTCCAAACAAGCGATGGAATTTGAAGAGGCTCTGGTGGCCCATCTGTGTACTTGTG	778
Db	696	GTCTGGAAGATCTCAGATGGCATCCA---GGACCTGGGGTCCCTGCGCTGGGAGTGGTC	752
OY	779	TTGTGTCTATTCATTCGATGGCTCATGTTGTTGGAGTCTGCGCCGAGAGTCAAGAT	838
Db	753	CTGTGCTCCGCTGGTGGCTGGATCATCTGATTTTCTGCAATCTGGAAAAGGGGTCAAGTCC	812
OY	839	TCCGGCAAGCGGCTATCTCTCGCGCTTCCCATACGTTGTCATATCATCTTATTC	898
Db	813	ACAAGCAAGTGGTGTACTTCACAGCTACTTCCCTTACGTCATGCTGTGTCTGTTG	872
OY	899	ATCACACATCATCTCCCGGTGTCTACTAGCAGGCACTCTGTTCTTCTGTCAGGCTCAA	958
Db	873	ATCCGAGAGATMACACTGCTGGAGACGCCACGGAATTCAGTTTACTCTTACCCCAAC	932
OY	959	TGGGCAAACTCTTGAGCTGGGTATGATACAGACGAGTACGCAAGTGTCTCTCT	1018
Db	933	ATCACAGCTCTGTGGGATCCCAAGTGTGATGATGTGGGGACCCCAATCTCTTCTCC	992
OY	1019	CTGACAGTGCACCCGACCGCATCATCTATCTCTTCAACAGGTTTGACATCAAT	1078
Db	993	TTTACCATCTCTGGGGGTGCTCAGCGCCCTGGGCACTCAACAAATACACAAACAAC	1052
OY	1079	ATTCACAGGAATGCTTGATTTGTTACGACTTTGGACACTTTACAAGTTTCTGTCTGG	1138
Db	1053	TGCTACAGGGAGCTGCGTCCCTTGTGCAATTCACAACACACACACACAGCTTGTGGCCGG	1112
OY	1139	TGCACGATCTTCGGTATCTTGTAACCTCCGCTACGAACTCAACTCAGAGTGGAGAT	1199
Db	1113	TTTGCATCTCTCCATCTGTGGGCTATATGCTCAGAGGACAGGGGTATCCAT---ATCT	1165
OY	1199	GTGTGCGTCTGGCGGTACAGTCTTGCTTCAATTTCAATACCTGATGCCATTGCAAA	1258
Db	1170	GAGGTGTGTGATCAGGCGCTGCGCTGCTATTCATCCGCTACCTCGAGCTGTGGTATG	1228
OY	1259	ACATTCACAACCTCAGCTATCTCGGTGCTGTTCTTCGAGAGTGTGGTCTGGGTATC	1318
Db	1230	TTACCTTCTCGCCTTTGTGGGCTGCTGTTTCTTCTTCAAGTGTGTTCTCTGGAGCTA	1288

OY 1339 GGCATCCGGGCTCTCTTCATCCACTTTCAACCATTTGGGGATGGAGCGCTTCCCAAGT 1378  
 Db 1290 GACAGCCAGTTTGTGTGTGTGTGAGAAAGCCTCGTACAGCGCTGTGTGGACATATATCCCGG 1349  
 OY 1379 GTACCACCGCTCTACATGTGACGATGACC-----TGTCTTGGGGTTTC 1423  
 Db 1350 GTGTTCCGTAAAGAAAGAACGGAGGAAATTCATCTCATCGTGTGTGTGCTCTTTC 1409  
 OY 1424 CTGCTTGGACTTGTTTACTGCACACCGGGTGGACAAATATATTCTTGACCTTGTAGATAC 1483  
 Db 1410 TTCAATCGGGCTCAATATAGCTCTACAGAGGGCGGAGTACGTTCCAGCTCTTGACATAC 1469  
 OY 1484 TACGGTGAAATTCCTCTTGCTTTTTCG---GCCATTTGTGAACTCGAGGGGTGTC 1540  
 Db 1470 TATGCGGCAAGTGGCATGTCTTCTCTTTTGTGGCATTGTAGTCCCTCTGTGTGGCT 1529  
 OY 1541 TGGATTTATGGATTGGAGAAATCTGTGCCATAGACATTTGATTCATTTGGTAAAGACT 1600  
 Db 1530 TGGGTTTACGGAGCGACGCCGCTTCTATGACAAACATTTGAAGATATGATTGGGTACAGCGG 1589  
 OY 1601 GGTGTTTACTGGCGCTCTGCTGTGGGGCTAATACACTCTGCTATTAATGACACTGTGTTT 1666  
 Db 1590 TGGCCTCTTAACAATACCTGTGGCTCTTTTTCACGGCAGCGTGTGGCTTGGCAACCTTC 1649  
 OY 1661 TTCTACGGCTCTTCCGCTCTTAACACACTGGGTTCGGAGACAACATAGTAATACC 1716  
 Db 1650 CTGTTCTCCCTGATCAATATACAGCGCACTGACCTTAACAAGAAAGTACACATATCC 1705

```

RESULT 12
US-09-843-598-1
: Sequence 1, Application US/09843598
: Patent No. US20020010944A1
:
GENERAL INFORMATION:
: APPLICANT: Horvitz, H. Robert
: APPLICANT: Ranganathan, Rajesh
: TITLE OF INVENTION: CESTR GENES, PROTEINS, AND MODULATORY
: TITLE OF INVENTION: COMPOUNDS
: FILE REFERENCE: 01997/525002
: CURRENT APPLICATION NUMBER: US/09/843,598
: CURRENT FILING DATE: 2001-04-26
: PRIOR APPLICATION NUMBER: US 60/200,549
: PRIOR FILING DATE: 2000-04-26
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 2016
: TYPE: DNA
: ORGANISM: Caenorhabditis elegans
US-09-843-598-1

```

Query Match	5.1%	Score 146.8	DB 10	Length 2016
Best Local Similarity	46.3%	Pred. 2.1e-21		
Matches 644	Conservative 0	Mismatches 722	Indels 24	Gaps 4
QY	259	ATGTCACAACATTTGAATTCCTTGATGTCCTGCATCGCTACATCCGCTGGTTGGGTAA	318	
Db	306	ATGGCAACTAAATGGAATTCCTGTTGGCCGCTGGTGGATATGCAATTTGGGTAA	365	
QY	319	CGTGGGCGGTTCCCTTTCATCGCGTACCAAGATGGAGAGTGCTTCCGTGGCCATA	378	
Db	366	TATAAGGGGAATCCCATCATGATGCTACAAAACAGGTGGCGCTGTTCTTAATTCATA	425	
QY	379	CGTATCGTCTTTTACTGTGCGGCAACCTGTGACTACTTAGAGTGTGCTCCGGACA	438	
Db	426	TTTCATTATGTTAATGATCGAGGACCTCCCATGTTCTATATGGACTTGTACTCGGACA	485	
QY	439	ATTCAGTTCAGAAGAACTCTGTTAAAGTTGGTGCAA--TTTCAACCGGCATGAAGAGTAC	495	
Db	486	ATTTCATCGGTACAGAGATGTGTTAGATATGGAAGAGTGTCGCCGTGTTCCAGGAAT	545	
QY	496	TGATACGCTCAAGCTGCGGCGCTGGGTTACATCCTGTCTTACAGTGGTGATCTGTGG	555	

Db 546 CGGTTACGGTATCTGCTATTGTCAGCTTCAATAGCCATTTTCTATATGCAATCCG 605  
QY 556 TC-----TCTGCTGATTTACTAGTATGAGCTTCGAGCCACTTCCTCAG 603  
Db 606 TCAAGCCGTCTATTGCTATTGTTGACCTTTCAAAAATTGGGATTCGAAAGTTCCG 665  
QY 604 GGCATTATTTGACGCTGATGGAGAACTGCTACCTGATATCCAACTTGCTGCATC 663  
Db 666 GGCCTCATTTGGCAATCCGCTGGAAATACACGAGATCTCAGATGACCTCAGTCAAT 725  
QY 664 AGTCAACAATCACCATATGATGACAGAGCTCACTTCACTTTTGAAGACAGTCT 723  
Db 726 ATCTGAAATATGGAGACACCATTTGACCTCCGTCAGAGAAATATATTATACAACTCT 785  
QY 724 CCAA---CAAGGATGGAATGGAAGAGTCTCGGTGCCCATCTGCTACTGCTGT 780  
Db 786 TGAAGTTCAAAAATCAACAGATTCGATATCTTGGAGGTGAATAAATCTCAATGGCACT 845  
QY 781 GTGCTATTATCATGATGCTCATGCTGATGCTGAGTCCGCGCCGAGATCAAGAGTTC 840  
Db 846 GTGCTACTGCTGATTTATATATGTTTACTTTCCTTTGGAAGGCTCCACAGTCTC 905  
QY 841 CGGCAAAAGCGCTTACTTCTGCGCTCTTCCCATACGTTGTCATGATCATTTATCAT 900  
Db 906 TGGAAAAATTTGTTGGGTACTGCAACAGCTCCATATATTATCTAAGTATTTCTCTAT 965  
QY 901 CACCAACATCATCTGCGCGGTCTACTGACGGCATCTGTTCTCTGCTACCGCTCAAG 960  
Db 966 AGGTGACTTCTTCTTCTGAGACAAAGATGCTCTTATTTATGATGACACCGGATTT 1025  
QY 961 GCGCAAACTCTTGGAGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1020  
Db 1026 CAGAAACTCAAGGATCTGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1085  
QY 1021 GACAGTGTACACGGACCATCATCATCTTCTCTTCAACAGGTTTACAGATTAAT 1080  
Db 1086 TGGACAGGATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1145  
QY 1081 CTACAGGATGCTTGGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
Db 1146 CATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1205  
QY 1141 CACGATCTTGGTATCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
Db 1206 TGTGATATCTTACACTTGGCTATGCTCTTCTCACCAATTAACCATTAATAGAT 1265  
QY 1201 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
Db 1266 AGTTGAGAACAGAGAGCTCTCTATCTTCACTGCTACCCCAAGCCCTCGACATAT 1325  
QY 1261 ATTCAACCTGACTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
Db 1326 GATTTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1385  
QY 1321 CTATATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
Db 1386 CTTCCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1445  
QY 1381 ACCCAACCTGACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1434  
Db 1446 GTGCAAAAATTCGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1505  
QY 1435 TGTATTACTGACAGCGGTGACAAATATATCTTGAAGTGTGATGATGATGATGATGAT 1494  
Db 1506 TCCCGCTATACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1565  
QY 1495 ATTCTTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1554  
Db 1566 TCTATCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1625  
QY 1555 GAGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1614  
Db 1626 TATATGATTTCAAAAAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1685

QY 1615 TCTCTGCTGG 1624  
Db 1686 AGTCTGCTGG 1695

RESULT 13  
US-09-843-598-3  
; Sequence 3, Application US/09843598  
; Patent No. US2002010944A1  
; GENERAL INFORMATION:  
; APPLICANT: Horvitz, H. Robert  
; APPLICANT: Ranganathan, Rajesh  
; TITLE OF INVENTION: CESPERT GENES, PROTEINS, AND MODULATORY  
; FILE REFERENCE: 01997/525002  
; CURRENT APPLICATION NUMBER: US/09/843,598  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR FILING DATE: 2000-04-26  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 2016  
; TYPE: DNA  
; ORGANISM: Caenorhabditis elegans  
US-09-843-598-3

Query Match 5.1%; Score 146.8; DB 10; Length 2016;  
Best Local Similarity 46.3%; Pred. No. 2.1e-21;  
Matches 644; Conservative 0; Mismatches 722; Indels 24; Gaps 4;

QY 259 ATGGTCCACAACTGATGATTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 318  
Db 306 ATGGGCACTTAATAAGGATTTCTGTTGGCGCTGTTGATATGCAATGATTTGGGTAA 365  
QY 319 CGTGTGGCGGTTCCCTTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 378  
Db 366 TATATGCGATTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 425  
QY 379 CGTATCGTCTTTTACTTGTGGCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 438  
Db 426 TTTCAATTAATTAATGATGAGGACTTCCATGTTCTATATGAACTTGTACTGAGACA 485  
QY 439 ATTGATTTCAAAATCTGTTAAAGTTGCTCAAA---TTTCAACCGCATTAAGATGATC 495  
Db 486 ATTTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 545  
QY 496 TGGATAGCTCAAGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 555  
Db 546 CGGTTACGGTATCTGCTGATTTTGGACGCTTCAATAGCCATTTCTATTAATGCGATATCCG 605  
QY 556 TC-----TCTGCTGATTTACTAGTATGAGCTTCGAGCCACTTCCTCAG 603  
Db 606 TCAAGCCGTCTATTGCTATTGTTGACTTCTTCAAAAATTTGGGATTCGAAAGTTCG 665  
QY 604 GGCATTATTTGACGCTGATGGAGAACTGCTACCTGATATCCAACTTGCTGCATC 663  
Db 666 GGCCTCATTTGGCAATCCGCTGGAAATACACGAGATCTCAGATGACCTCAGTCAAT 725  
QY 664 AGTCAACAATCACCATATGATGACAGAGCTCACTTCACTTTTGAAGACAGTCT 723  
Db 726 ATCTGAAATATGGAGACACCATTTGACCTCCGTCAGAGAAATATATTATTAACAAAGTCT 785  
QY 724 CCAA---CAAGGATGGAATGGAAGAGTCTCGGTGCCCATCTGCTACTGCTGT 780  
Db 786 TGAAGTTCAAAAATCAACAGATTCGATATCTTGGAGGTGAATAAATCTCAATGGCACT 845  
QY 781 GTGCTATTATCATGATGCTCATGCTGATGCTGAGTCCGCGCCGAGATCAAGAGTTC 840  
Db 846 GTGCTACTGCTGATTTATATATGTTTACTTTCCTTTGGAAGGCTCCACAGTCTC 905  
QY 841 CGGCAAAAGCGCTTACTTCTGCGCTCTTCCCATACGTTGTCATGATCATTTATTCAT 900









DR WPI: 2001-441879/47.  
 DR N-PSDB; AAD09681.  
 XX  
 PT Novel invertebrate symporter cell surface receptor proteins and nucleic  
 PT acid encoding the protein useful as pesticide or drug target and to  
 PT identify compounds that have utility as pesticides or pesticides.  
 XX  
 PS Claim 18; Page 62-63; 71pp; English.

XX The invention relates to invertebrate symporter cell surface receptors  
 CC of the sodium/neurotransmitter family (SNF) and nucleic acid molecules  
 CC encoding such receptors. The SNF protein is useful for detecting a  
 CC candidate compound especially a putative pesticidal or pharmaceutical  
 CC agent that interacts with an invertebrate symporter cell surface  
 CC receptor protein or its fragment. Insect or worm genetically modified  
 CC to express protein of the invention is useful for studying invertebrate  
 CC symporter cell surface receptor protein activity by detecting the  
 CC phenotype caused by the expression or mis-expression of the protein in  
 CC the animal. Nucleic acids encoding the invertebrate receptor protein or  
 CC their fragments are useful as biopesticides. SNF nucleic acids are  
 CC useful for generating mutant phenotypes in an animal model or living  
 CC cells that are used to study the regulation of genes encoding the  
 CC proteins which are useful as pesticide or drug targets. The genetically  
 CC modified organisms or cells are useful in screening assays to identify of  
 CC pesticides or therapeutics and thus are useful in the identification of  
 CC new drug targets, therapeutic agents, diagnostics and prognostics  
 CC useful in treatment of disorders associated with ion channels. The  
 CC nucleic acid molecules are also useful as hybridisation probes.  
 CC The present sequence is *Drosophila melanogaster* (dm) K+ coupled  
 CC amino acid transporter of the SNF family which is referred as dmSNF.  
 CC  
 XX Sequence 641 AA:

Query Match 47.0%; Score 1577.5; DB 22; Length 641;  
 Best Local Similarity 50.5%; Pred. No. 4e-156;

Matches 314; Conservative 90; Mismatches 183; Indels 35; Gaps 9;

QY 30 ANKALDNDITDLEAEPERWWSNIEFLMSCIATSVGLGNVWFPEIAYONGGAP 89  
 DB 20 ATNAASSTK---TDAEPTAERTNMGNGLEFLMSCISVSGLNVMFPPTAYENGCAF 76  
 QY 90 LVPYVYVLLVGRKPYVLECVLGOFSSRNKYKMS:SPARKGKYQAACGTYLSYVY 149  
 DB 77 LIRYIYVFLIGRPMTYLEMIMKGFISQTKIMSVPGVGVGAFETICITISYSS 136  
 QY 150 ICGLCYLYLAMSFOATLPMAICQPEWENCVPSPDTLAASVNNITNGT----- 196  
 DB 137 LIALTYLYLFSFQSELPMSYCRDEWNTCVNSRP--QEVYDNLITGSLANESARNLSGI 194  
 QY 197 -----SSAQLYFLRTVLOOSDIEGGLCAPTYVLYLCFIAMLVFGYVARGYS 246  
 DB 195 VANDETERKLOSSSELYFLNVIYIEKDISDVGDPKMKTLALFVAMVYFVILIMRGYS 254  
 QY 247 SGRAAVFLALFPYVAMITLITITLPGATDGIFFVTYPMQAKLIEGVYSAVTOVFFS 306  
 DB 255 SGRAAVFLALFPYVLEVLIRAVTLEGARADGILFLEPQMGELANPWTKEAVQCFES 314  
 QY 307 LVTCTPPIIMESSYNGFRNHIYDAMIYVTLDTFTSFLSGCTIFGLIGNAYELNSE-VG 365  
 DB 315 LAVSGSPIIMFASYNFHDGIYRDAMIYVTLDTLSTLGGITITFALIGNAHMLQIENIR 374  
 QY 366 DVGAGGTSIAFSYSDATAKTRQ--POLFSVLEFFLMASVLTGIGSSVALLSTNTLAMD 423  
 DB 375 DVV-RSGTGLATISYDAISK--FOAVPOLFSVLEFFMLFVLGIGSIVALLQSTIVTTICQ 432  
 QY 424 FPRVPTVYAMTSCGFLGLGYCTPGGQYILELDHYHGGFLVFCALISALFEMWY 483  
 DB 433 FKGWKYKVALTITVSGFLMGVLYVTPGGQMLITLVDGTYVVFILAFELAGLVWY 492  
 QY 484 GLENLCLDIEMLGKTKTGAWRKCGVITPAIMTVFYALLASNNLVFGDNVYPTAGY 543  
 DB 493 GLONFCDIEFMGNRRVSLYWRVCMSEFTPVMMIIFISYMTVIEIKYSELY-FPEAN 551

QY 544 VSGYLMFLGMTFVPIDIGISLYKRTGSEITIKAFHSKPSWSPRREREMQFKA 603  
 DB 552 IAGWLFPAIGAOPPLMGWLTYSRHPGTGYWKSLSKASLPDSRWPANPEIRREWYLFKN 611  
 QY 604 EAKALR-QKMNTSRVKHLWYSI 624  
 DB 612 QKAAGRATQKDTSLKGFWMKV 633

# RESULT 2

ABBS9760  
 ID ABBS9760 standard; Protein; 593 AA.

XX  
 AC ABBS9760;  
 XX

DT 26-MAR-2002 (first entry)  
 XX

DE *Drosophila melanogaster* polypeptide SEQ ID NO 6072.  
 XX

KW *Drosophila*: developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX

OS *Drosophila melanogaster*.  
 XX

PN W0200171042-A2.  
 XX

PD 27-SEP-2001.  
 XX

PF 23-MAR-2001; 2001WO-US09231.  
 XX

PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.

XX  
 PA (PEKE ) PE CORP NY.

XX  
 PT Venter JC, Adams M, Li PWD, Myers EW;  
 XX

XX  
 DR WPI: 2001-656860/75.  
 DR N-PSDB; ABL03863.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX

PS Disclosure; SEQ ID NO 6072; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB101840-AB116175) and the encoded proteins  
 CC sequences (ABBS7737-ABBS72072).  
 CC (ABBS7737-ABBS72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pcl\_sequences.

XX  
 SQ Sequence 593 AA;

Query Match 45.4%; Score 1523; DB 22; Length 593;  
 Best Local Similarity 51.0%; Pred. No. 1.9e-150;  
 Matches 301; Conservative 87; Mismatches 170; Indels 32; Gaps 8;

QY 62 MSCIATSVGLGNVWRFPIYONGGAFVYVYVLLVGRKPYVLECVLGOFSSRNYSK 121  
 DB 1 MSCIATSVGLGNVWRFPIYONGGAFVYVYVLLVGRKPYVLECVLGOFSSRNYSK 121

QY 122 VMSISPMKGTGYAAGCGYIISYVVICGLCLVYLYLAMSFOATLPMAICQPEWENCVPS 181  
 DB 61 IMSVGVGFGVGVGAGFGLICISYSSLLALTYLVFSGFELPMSYCRDEWNTCVNS 120

QY 182 DPTLAASVNNITNGT-----SSAQLYFLRTVLOOSDIEGGL 218

```

Db      121 RP--OEYDNLTLTGSILANESARNLSGIYANDETEKLOSSSELYELNVIKEKIDISDGV 178
OY      219 GAPIVYLICFIAMLWEGVARGVSSGKAAYFLAPYVMTLTFTTILPGADG 278
Db      179 GDPOMKLTALFVAVVILFVIMKGVSSGKAAYFLAPYVLTILIRATTLGADG 238
OY      279 ILFEVTPQAKLLELGVVYSAVTOVFSLVTCGPIIMFSSYNGFRNHYRDAMVITLD 338
Db      229 ILFELEPQMGELLNPTWKAEAVQCFSLAVSGPIIMFASYNRRDHGIRNAMLVTLLD 298
OY      339 TETSLSCSTIFGILGNATYELNSE-VGDVYAGAGSTLAFISYPAIAKTPO--POLFSV 395
Db      299 TLTSILGGITTFALIGNLAHNLQENIRDV- RSGTGLAFISYPAISK-FQAVPOLFSV 356
OY      396 LFFLLMVLIGISSVALSTFNTLMDAFPRVPTVYMSAMTSCGFLGLVITCGGGGY 455
Db      357 LFFFLMVLIGISSVALSTFNTLMDAFPRVPTVYMSAMTSCGFLGLVITCGGGGY 416
OY      456 LELVDHYGTEFLVFCALSELAVGVIMYLENCLDIEFMLGKKTGAYWRCLMGVITPAI 515
Db      417 LTLVDFYGGTYVVFILALFELAGIYVWYGLQNFCDIEFMCNRVSLVWRCVMSFETPYM 476
OY      516 MTTVFYFALLASNNLVFGDNVYPTAGVYSGTLMFLGNTFPIGIGSLYKRTGTSE 575
Db      477 MIIETYSMTYETPIKYSLEY-FPEANINAGWLFALGAOPPLMGWYISHHPOGTYWK 535
OY      576 TIKKAFHSKPSWGPSPRREREMOFAKAKALR-OKMNTSRVKKLWYSI 624
Db      536 SIKASLKSDFDRGPNPEIRREWVIFKNOKAQRATQDTSKLGFEFWKV 585

```

# RESULT 3 ABB58403

ID ABB58403 standard; Protein; 651 AA.

AC ABB58403;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 2001.

KW Drosophila: developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.

PN 27-SEP-2001.

PD 23-MAR-2001; 2001WO-US09231.

PF 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

PA Venter JC, Adams M, Li PMD, Myers EM;

PI WPI; 2001-656860/75.

XX N-PSDB; ABL02506.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

PS Disclosure; SEQ ID NO 2001; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of

CC

CC insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB16176-AB130511), expressed DNA sequences (AB101840-AB16175) and the encoded proteins (AB57737-AB572072).

CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 651 AA:

Query Match 44.5%; Score 1492; DB 22; Length 651;

Best Local Similarity 48.4%; Pred. No. 3.9e-147;

Matches 305; Conservative 96; Mismatches 191; Indels 38; Gaps 12;

OY 11 ESSEPKMEPRSSQSLTPPANKKALDNIDTDLBAEPP--ERWVSNIEPIAMCIATYS 68

Db 38 ESSNKKTLSESGSAPRPHGDS-----PQGVNAGQPKRRDWMNDIERLMSCIALS 91

OY 69 VGLGNVWRPPTIAYNGGAFVPPYVYLLVGLKRPVYLYECVLGQFSSRNYSKWSISPA 128

Db 92 VGLGNVWRPPTIAYNGGAFVPPYVYLLVGLKRPVYLYECVLGQFSSRNYSKWSISPA 151

OY 129 MKGTGTAAAGGTYLLSYVVICGLCYLYLAASPDATLPWATCOPEM-ENCYPSDPTLAA 187

Db 152 MKGIGYGOVLATGYVTTYATLMLALTLRFDVDFPTLPWYSYCREMGTECLDSGPQKAS 211

OY 188 SVNNITNGTSSAQLVFLRTLOQSDIGELGAPITWYVLCFIAMLWEGVARGVSS 247

Db 212 RATSLAGCGGTHSPFRNRIILREKASIDIGIYPSKSLALAAVAVIYAGIMFGVSS 271

OY 248 GKAAYFLALFPYVMTLTFTTILPGATDGLFEVTPQAKLLELGVVYSAVTOVFESL 307

Db 272 GKASYFLALFPYVMTLTFTTILPGADGVLFEFRPMKHLLEPQVWYAAVTOVFESL 331

OY 308 TYCTGPIIMFSSYNGFRNHYRDAMVITLDTFTSLSCSTIFGILGNATYELN-SEVGD 366

Db 332 AICFGNIIWYASYNRGNHYRDAMVITLDTFTSLSCSTIFGILGNATYELNNTDIALS 391

OY 367 VVGAGGTSIAFISYPAIAIK-TEPOLFSLVFLMMSVIGISSVALLSTFNTLMDAF 425

Db 392 VVN-GGPGIAFISYDAIAIKFKMLQLREVEFLMFLVIGISSNMGMSCTVAKDQG 450

OY 426 RVPYVMSAMTSCGFLGLVYCTPGQYILEVDHYGTEFLVFCALSELAVGFWIYGL 485

Db 451 HLKMTVYVVGIAIVGFLGLVITPGQGLMLNVDGVTVALVLAIRVLTAMIVG 510

OY 486 ENLCIDIEPMLGKKTGAYWRCLMGVITPAIMTTFYFALLASNNLVFGNYYPTAGIYS 545

Db 511 KRLCRDVERMIGIKTSLYRICWAVVTPMLMTLITLYLVLEPKYKD-YTYQSGVYVF 569

OY 546 GTMLFLGMTFVPDYGIFSLY-----KYRTGTFSEETIKKAFHSKPSMGPRSPRR 596

Db 570 GWCL-----SAFGVGVLFMAIPAVRKQPSHLGLMAR-IRKAFEPPLMGPSPDPTLK 621

OY 597 EMMQFKAE--AKALRQKNTSRVKKLWYSI 624

Db 622 RYOLFVQEGGNANALFRSS-----IMWKI 645

XX RESULT 4

XX ABB71178

XX ID ABB71178 standard; Protein; 639 AA.

XX ABB71178;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 40326.

XX Drosophila: developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.  
 XX 27-SEP-2001.  
 XX 23-MAR-2001; 2001WO-US09231.  
 XX 23-MAR-2000; 2000US-191637P.  
 XX 11-JUL-2000; 2000US-0614150.  
 XX (PEKE ) PE CORP NY.  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 XX WPI: 2001-656860/75.  
 XX N-PSDB; ABL15281.  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 XX genes from Drosophila and for elucidating cell signalling and cell-cell  
 XX interactions -  
 XX  
 XX Disclosure: SEQ ID NO 40326; 21pp + Sequence Listing; English.  
 XX  
 XX The invention relates to an isolated nucleic acid detection reagent  
 XX capable of detecting 1000 or more genes from Drosophila. The invention is  
 XX useful in developmental biology and in elucidating cell signalling and  
 XX cell-cell interactions in higher eukaryotes for the development of  
 XX insecticides, therapeutics and pharmaceutical drugs. The invention  
 XX discloses genomic DNA sequences (AB101840-AB116175) and the encoded DNA  
 XX sequences (AB101840-AB116175) and the encoded proteins  
 XX (AB101840-AB116175).  
 XX The sequence data for this patent did not form part of the printed  
 XX specification, but was obtained in electronic format directly from WIPO  
 XX at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 639 AA:

Query Match 43.98; Score 1472.5; DB 22; Length 639;  
 Best Local Similarity 47.88; Pred. No. 4.2e-145;  
 Matches 302; Conservative 97; Mismatches 200; Indels 33; Gaps 11;

QY 21 RSSQISIPANNKAL-DNIDDDLEAPPERAVWSNNITFLMSCTATSGLGWAFPP 79  
 DB 19 RSTASTVEISTNSPDKNSDDQDAKVPERRATWKGVEFLMSCLAMSGLVNMFPP 78  
 QY 80 IAYONGGAGFLVYVYLLVGRPVYLLCYLGQFSRNSNVKWSISPAKKGTYAAG 139  
 DB 79 TALDNGGAGFLVYVYLLVGRPVYLLCYLGQFSRNSNVKWSISPAKKGTYAAG 138  
 QY 140 CGYLLSYVYVYLLVGRPVYLLCYLGQFSRNSNVKWSISPAKKGTYAAG 192  
 DB 139 VFMLSTYVYVYLLVGRPVYLLCYLGQFSRNSNVKWSISPAKKGTYAAG 198  
 QY 193 TNGT-----SSAQLVFLRVYVYLLVGRPVYLLCYLGQFSRNSNVKWSISPAKKGTYAAG 243  
 DB 199 QNTTMSQNDRVYVYLLVGRPVYLLCYLGQFSRNSNVKWSISPAKKGTYAAG 258  
 QY 244 VKSSGAAVYVYLLVGRPVYLLCYLGQFSRNSNVKWSISPAKKGTYAAG 303  
 DB 259 VKSSGAAVYVYLLVGRPVYLLCYLGQFSRNSNVKWSISPAKKGTYAAG 318  
 QY 304 FSLVCTGPIVYVYLLVGRPVYLLCYLGQFSRNSNVKWSISPAKKGTYAAG 362  
 DB 319 FSLVCTGPIVYVYLLVGRPVYLLCYLGQFSRNSNVKWSISPAKKGTYAAG 378  
 QY 363 EGVGVAGAGTSLVYVYLLVGRPVYLLCYLGQFSRNSNVKWSISPAKKGTYAAG 420  
 DB 379 DIGSVV-KGGAGIAFSTIPDAIK-FKNLPQITSVYVYLLVGRPVYLLCYLGQFSRNSNVKWSISPAKKGTYAAG 436  
 QY 421 MDAPRPVYVYLLVGRPVYLLCYLGQFSRNSNVKWSISPAKKGTYAAG 480  
 DB 437 RDRFPVYVYLLVGRPVYLLCYLGQFSRNSNVKWSISPAKKGTYAAG 496

QY 481 WYGLNCLDIEFLLGKGTGAYWRLCGVITPAIMTVYFAYALLASNNLVEGDNVYPT 540  
 DB 497 WYGLNCLDIEFLLGKGTGAYWRLCGVITPAIMTVYFAYALLASNNLVEGDNVYPT 555  
 QY 541 AGVSGYLMFLGTFVPIGIFGLVYVYLLVGRPVYLLCYLGQFSRNSNVKWSISPAKKGTYAAG 600  
 DB 556 WYGLNCLDIEFLLGKGTGAYWRLCGVITPAIMTVYFAYALLASNNLVEGDNVYPT 613  
 QY 601 FKAFAKALROKMTSRKHLWYVYLLVGRPVYLLCYLGQFSRNSNVKWSISPAKKGTYAAG 632  
 DB 614 ---HKIEENELTPRKOGIW---AAIKONI 637

RESULT 5  
 AAE05104  
 ID AAE05104 standard; Protein: 639 AA.  
 XX AAE05104;  
 XX 18-SEP-2001 (first entry)  
 XX  
 XX Drosophila melanogaster dmsNF3.  
 XX  
 XX Fruit fly: Invertebrate symporter cell surface receptor; dmsNF3;  
 XX sodium/neurotransmitter family; SNF; biopesticide; therapy.  
 XX Drosophila melanogaster.  
 XX WO200149848-A2.  
 XX 12-JUL-2001.  
 XX 28-DEC-2000; 2000WO-US3551.  
 XX 30-DEC-1999; 99US-0173929.  
 XX 15-MAR-2000; 2000US-0189399.  
 XX 23-MAR-2000; 2000US-0191686.  
 XX 23-MAR-2000; 2000US-0191687.  
 XX 23-MAR-2000; 2000US-0191688.  
 XX 23-MAR-2000; 2000US-0191695.  
 XX (GENO-) GENOPTERA LLC.  
 XX Kellerman KA, Keegan KP, Edens AJ, Torpey J;  
 XX WPI: 2001-441879/47.  
 XX N-PSDB; AAD09684.  
 XX  
 XX Novel invertebrate symporter cell surface receptor proteins and nucleic  
 XX acid encoding the protein useful as pesticide or drug target and to  
 XX identify compounds that have utility as therapeutics or pesticides -  
 XX  
 XX Claim 18; Page 67-68; 71pp; English.  
 XX  
 XX The invention relates to invertebrate symporter cell surface receptors  
 XX of the sodium/neurotransmitter family (SNF) and nucleic acid molecules  
 XX encoding such receptors. The SNF protein is useful for detecting a  
 XX candidate compound especially a putative pesticidal or pharmaceutical  
 XX agent that interacts with an invertebrate symporter cell surface  
 XX receptor protein or its fragment. Insect or worm genetically modified  
 XX to express protein of the invention is useful for studying invertebrate  
 XX symporter cell surface receptor protein activity, by detecting the  
 XX phenotype caused by the expression or mis-expression of the protein in  
 XX the animal. Nucleic acids encoding the invertebrate receptor protein or  
 XX their fragments are useful as biopesticides. SNF nucleic acids are  
 XX useful for generating mutant phenotypes in an animal model or living  
 XX cells that are used to study the regulation of genes encoding the  
 XX proteins which are useful as pesticide or drug targets. The genetically  
 XX modified organisms or cells are useful in screening assays to identify  
 XX pesticides or therapeutics and thus are useful in the identification of  
 XX new drug targets, therapeutic agents, diagnostics and prognostics  
 XX useful in treatment of disorders associated with ion channels. The  
 XX nucleic acid molecules are also useful as hybridisation probes.

CC The present sequence is Drosophila melanogaster (dm) SNF homologue  
CC which is referred as dmsnf3.  
SQ Sequence 639 AA;

Query Match 43.9%; Score 1472.5; DB 22; Length 639;  
Best Local Similarity 47.8%; Pred. No. 4.2e-145;  
Matches 302; Conservative 97; Mismatches 200; Indels 33; Gaps 11;

```
OY 21 RSSQISLPPANNKAL-DNIDDTDLLEAPPERMWSNNIEFLMGSCLATSVGLGNWRPPE 79
DB 19 RSTASTVEISTNSPLRDNSDDQEAAYPEERATMGKVEFLMGLSCIASVGLGNWRPPE 78
OY 80 IAYNGGGAFLPYVIVLLVGVKPPVYLLVGVGFSSSNYSKVSISIPAMKGTGAQAAG 139
DB 79 TALDNGGGAFLPYVIVLLVGVKPPVYLLVGVGFSSSNYSKVSISIPAMKGTGAQAAG 138
OY 140 CGYLLSYVYVIGCLVLYLAMSFOATLFAICQPEWE-NCVPSD-----TLASVNNI 192
DB 139 VFMLSTYAAALVATIGRFIESFRNPFLMSTCRAMWGCHINSAPDASNMSQLESNDQRP 198
OY 193 TNGT-----SSAQLYFLRTVLOQSDGIEGGGAPIMVLYLCLFTAMLMVGVARG 243
DB 199 QNTYMKSONDRVITTSSEWYFVKEVLEHEKPNIEEGIGLPMWELVIGLFTAMSCVFFIRRG 258
OY 244 VKSSGKAAYFLALPEPYVMITLFTTIIIPGATDGIFFVYPOWAKILEGWSAVYQV 303
DB 259 VKSSGKAAYFLALPEPYVMITLFTTIIIPGATDGIFFVYPOWAKILEGWSAVYQV 303
OY 304 FFSLTCTGPIIMSSYNGFRNNTYRDAMITTTIDTFPSISGCTIFGILCNLAELNS- 362
DB 319 FYSLSVCEGNITIMSSFNKFGHNVHRDAIYTGDTMTSLAGFTIFGILHAEICTD 378
OY 363 EVGDVVGAGTSLAFISYPAIAKTFQ--POLFSVLEFLAMSVLGISGVALLSTFNTLA 420
DB 379 DIGSVY-KGAGAGLAFISTPDAIAK-FKNLPQISVLEFLMFLVIGSINIMTSCSVTAI 436
OY 421 MDAPRPVPTVYMSAMTSCGFLGLGYCTPGGOVYLLVDHYGTFVLVFCALISLAVF 480
DB 437 RDRPFPNGQMCSSLIAVVSFFIGLMTTPGGQVWLLVDFGASMLALVIGIALVLTIG 496
OY 481 WYIGLNLCDIEFLMGLKKTGAWRLCMGVTPPAIMTTFVFFYALLASNNLVFGDNYVPT 540
DB 497 WYIGDRLCKDIEFLMGLKKTGAWRLCMGVTPPAIMTTFVFFYALLASNNLVFGDNYVPT 540
OY 541 AGYSGVYMLFLMGTFVPIGIGSLKYKRTGTFSEFIKKAHRSKMGPSRPREREM 600
DB 556 WSTIGMLITAFGLIQLPIMIVAVLRDPPGTLGAKINGAFTPKKMGPSDPLREQY-- 613
OY 601 FKAERKALRQMNTSRVKNHLYWSTIGAYRNI 632
DB 614 ----HKEIENELTPRGOGIW----AAIKONI 637
```

RESULT 6  
ABB64027  
ID ABB64027 standard; Protein; 629 AA.  
AC ABB64027;  
DT 26-MAR-2002 (first entry)  
XX Drosophila melanogaster polypeptide SEQ ID NO 18873.  
DE Drosophila melanogaster polypeptide SEQ ID NO 18873.  
XX Drosophila melanogaster polypeptide SEQ ID NO 18873.  
KW Drosophila melanogaster polypeptide; cell signalling; insecticide;  
KW pharmaceutical.  
OS Drosophila melanogaster.  
XX WO200171042-A2.  
PN 27-SEP-2001.  
PD

PF 23-MAR-2001; 2001WO-US09231.  
XX 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX (PEKE ) PE CORP NY.  
PA Venter JC, Adams M, Li PWD, Myers EW;  
PI WPI: 2001-656860/75.  
XX N-PSDB; ABL08130.

DR WPI: 2001-656860/75.  
XX N-PSDB; ABL08130.  
DR New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions -  
PS Disclosure: SEQ ID NO 18873; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB57737-AB872072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcl\_sequences.

Sequence 629 AA:

Query Match 42.2%; Score 1414; DB 22; Length 629;  
Best Local Similarity 43.9%; Pred. No. 5.6e-139;  
Matches 280; Conservative 113; Mismatches 209; Indels 36; Gaps 9;

```
OY 10 FESSEPKMEPRSSQISLPPANNKALDNIDDTDLLEAPPE-----RMYMS 55
DB 7 YOKLRNNOGAOGSSEHGRGN-----DGISTVYISABGEELTINCEASESSGGRDQMS 61
OY 56 NNIPLMCAITSVGLGVNWRPPIAYONGGAPLVPIVYLLVGVKPPVYLLVGVGFSS 115
DB 62 RGVLEFLFSCIALSVGLGVNWRPPIAYONGGAPLVPIVYLLVGVKPPVYLLVGVGFSS 121
OY 116 SRNSYKVSISIPAMKGTGAQAAGGCTLYVYVIGCLVLYLAMSFOATLFAICQPEW 175
DB 122 SRGCTRAFDPAPIMNGIAYGOVYSTALATTYACIMALTIRLVASFSEVLPWYICLVEM 181
OY 176 -ENCVPSDPTLAASVNNITNGTSSAQLYFLRTVLOQSDGI-EGGIGAPIMVLYLCLFTAW 233
DB 182 GKSCVATGAT-AANDSSIVGVSSAELEFFGTQVLRPEPSLDNDNGIGTSPMDLVCLATW 240
OY 234 LMVGVVARGYKSSKAKAYFLALPPYVMITLFTTIIIPGATDGIFFVYPOWAKILEG 293
DB 241 VIIGTILSKGIRSSKASAYFLALPEYVIMIVLIRAVLPGAMGIVYELRPQMSQLNLP 300
OY 294 GWYSAVYOVFFSLVCTGPIIMFSSYNGFRNNTYRDAMITTTIDTFPSISGCTIFGIL 353
DB 301 HWYTAITQMFPSLAICGTLVMSYFDFPNKNNKVDYITTTIDSLTSLAGCTIFGIL 360
OY 354 GNLAVELNSEVDVAGGTSIAFISYPAIAK-TFOPQLFSVLEFLMGSVGLIGSSVAL 412
DB 361 GNLAFETWTKDISQVVGKAGIATSYDEALAFKRYLLQDLAVLEFFMLVLIGISNIGM 420
OY 413 LSTFNTLMDAFPRPYVYMSAMTSCGFLGLGYCTPGGOVYLLVDHYGTFVLVFCAL 472
DB 421 ASAVVNVYKDRFTHLPHWLLAVSASIIIGFLGLVYMTGQGVFLVLPFYCGTFLALVAL 480
OY 473 ISELAVFMVIGLEMLCDIEFLMGLKKTGAWRLCMGVTPPAIMTTFVFFYALLASNNLV 532
DB 481 IAEIYAVGMITGYKICSDIEFLMGLKKTGAWRLCMGVTPPAIMTTFVFFYALLASNNLV 540
OY 533 GDNVYVPTAGYVGYLMLFLMGTFVPIGIGSLKYKRTGTFSEFIKKAHRSKMGPSR 592
```

DB 541 -RGVQYPPVYVMGMIIMGLVQLPFWALVTIYQCGKTEGSGKFGIAMPOTAWGSPLOT 599  
QY 593 RERRENMOKRAEKALROK-----MNTSRVKHLMYSTIG 626  
DB 600 Q-----KFEAYIIHRKREADFKSPRGYLEDNIFG 629

RESULT 7  
AAE00297  
ID AAE00297 standard; Protein: 629 AA.  
AC AAE00297;  
XX  
DT 13-JUN-2001 (first entry)  
XX  
DE Drosophila melanogaster bioamine transporter 1 (Bt1) homologue.  
XX  
KW Fruitfly; transporter protein; bioamine transporter 1; Bt1; screening;  
KW therapy; molecular transport disorder; biopesticide;  
KW invertebrate model.  
XX  
OS Drosophila melanogaster.  
XX  
FH Key  
FT Domain  
FT 57..590  
FT /label= SNF domain  
FT /note= "Sodium neurotransmitter symporter family,  
FT specifically claimed in claim 1"  
FT 66..88  
FT /label= Transmembrane\_domain\_1  
FT 92..118  
FT /label= Transmembrane\_domain\_2  
FT 137..157  
FT /label= Transmembrane\_domain\_3  
FT 233..248  
FT /label= Transmembrane\_domain\_4  
FT 257..274  
FT /label= Transmembrane\_domain\_5  
FT 302..328  
FT /label= Transmembrane\_domain\_6  
FT 338..362  
FT /label= Transmembrane\_domain\_7  
FT 400..422  
FT /label= Transmembrane\_domain\_8  
FT 435..457  
FT /label= Transmembrane\_domain\_9  
FT 459..482  
FT /label= Transmembrane\_domain\_10  
FT 514..533  
FT /label= Transmembrane\_domain\_11  
FT 551..571  
FT /label= Transmembrane\_domain\_12  
XX  
PN WO200118178-A1.  
XX  
PD 15-MAR-2001.  
XX  
PF 08-SEP-2000; 2000WO-US24598.  
XX  
PR 09-SEP-1999; 99US-0153461.  
XX  
PA (GENO-) GENOPTERA LLC.  
XX  
PI Ebens AV, Keegan KP, Winslow JM;  
XX  
DR WPI: 2001-235196/24.  
DR N-PSDB; AAD03450.  
XX  
PT Drosophila melanogaster Bioamine Transporter 1 (Bt1) nucleic acid and  
PT protein, useful in screening assays to identify candidate compounds  
PT which are potential pesticide agents or therapeutics that interact with  
PT Bt1 proteins

XX  
PS Claim 2; Page 49-51; 53pp; English.  
XX  
CC The present sequence is a Drosophila melanogaster transporter protein  
CC homologue, bioamine transporter 1 (Bt1). The Bt1 is used in screening  
CC assays to identify candidate compounds which are potential pesticide  
CC agents or therapeutics that interact with Bt1 proteins. It can also be  
CC used to genetically modify metazoan invertebrate animals resulting in  
CC Bt1 expression or mis-expression. It can also be used in methods for  
CC identifying new drug targets, therapeutic agents, diagnostics and  
CC prognostics useful in the treatment of disorders associated with  
CC molecular transport across the membrane. The Bt1 nucleic acid or its  
CC fragments, such as an antisense sequence or double stranded RNA, may be  
CC used as a biopesticide to inhibit Bt1 function. The invertebrate model  
CC organisms such as Drosophila melanogaster are useful for rapidly carrying  
CC out large-scale systematic genetic screens and therefore for analysing  
CC expression and mis-expression of Bt1 protein.  
XX  
SQ Sequence 629 AA;  
XX  
Query Match 42.1%; Score 1413; DB 22; Length 629;  
Best Local Similarity 43.9%; Pred. No. 7.2e-139;  
Matches 280; Conservative 113; Mismatches 209; Indels 36; Gaps 9;  
QY 10 FESSEPKMEPRSSQISLPANNKKAALNDIDTDLAEPE-----RWWS 55  
DB 7 YOKLRMMQAGOSRREHPGSN-----DGISTYIYSRGEELTINCEASESSGORDWS 61  
QY 56 NNIEFLMSTIATSVGLGNVWRPFPIAYONGGAFVLPYVTVLLVNGRPVYELCVLGQFS 115  
DB 62 RGVEFLFSCIALSVGLGNVWRPFPIALENGGAFLLPYVTVLLIGRPVYELVITIGQFS 121  
QY 116 SRNSVYVMSISPMKGGVYVQAAGCCGIIISYVVICGLCYLXLMASQALPMAIQPEW 175  
DB 122 SRGCIKRFMDAPIMKRGVYVSTALATVYICMAITRYLVASEVLPPTYCLVW 181  
QY 176 -ENCVPDPTLAASVNNITNGSSAQLFLRTVLAQSDGI-EGSLGAPIVYLVLCETAW 233  
DB 182 GKSCVATGAT-AANDSIVGVSSAELEFQYTLREPSIDNDGLGTPSMDVLCLATW 240  
QY 234 LMFGVAVARGVSSGAAVFLAFPPVVMITLITITILPGANDGILFFTPPMAKITEL 293  
DB 241 VIIGTILSKIRSSGRASYFLALFPYIMIVLLIRAVTLPGAGQGYVFLKPMWSOLNP 300  
QY 294 GWNYSVATQVFSFLTQGPRIIMFSSYNGFRHNIYRDAMVTLDTFTSLSGCTFEJIL 353  
DB 301 HWYKAITQMFESLALCFGLVWYASFNDRKRVKNDVITITIDSLTSLACIIFGIL 360  
QY 354 GNLAYLELNSEVGDVAGGTSIAFISYDAIAK-TPQPLFSVLFPLMSVLGIGSSVAL 412  
DB 361 GNLAFTETNKDISQVYKGGAGLAFLISYDAIAKFKYLPQFANVLPFFMLLVGIGSNIGW 420  
QY 413 LSTFNTLADAFRRVFPVYVMSAMTSCSGFLGLGYCTPGQYITLEVDHKGTFVLVFECA 472  
DB 421 ASAIVNVVYDRFTHLPWMLAVASAIITGFCGLVYVMPGQFVLNVDFGCTFIALVA 480  
QY 473 ISELAVGFWIYGLNCLDIEFMLGKKTGAYVRCGVITPAITVFFYALASNNIYF 532  
DB 481 IAEILLAVGWITGYKRICSDIEFMLNKTSTYRICAIVAPGLMFLVLYVLMFSYEPLVY 540  
QY 533 GDNVYVPTAGVYSGVLMLEFGMTFVPIGIGFSLYKRYTGTFSETTIKAFHSKSPGRSP 592  
DB 541 -RGVQYPPVYVMGMIIMGLVQLPFWALVTIYQCGKTEGSGKFGIAMPOTAWGSPLOT 599  
QY 593 RERRENMOKRAEKALROK-----MNTSRVKHLMYSTIG 626  
DB 600 Q-----KFEAYIIHRKREADFKSPRGYLEDNIFG 629

RESULT 8  
ABB06092  
ID ABB06092 standard; Protein: 593 AA.  
XX

AC ABB60692;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 8868.  
 XX  
 KM Drosophila: developmental biology; cell signalling; insecticide;  
 XX pharmaceutical.  
 XX Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 XX 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI: 2001-656860/75.  
 XX N-PSDB; ABL04795.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 XX genes from Drosophila and for elucidating cell signalling and cell-cell  
 XX interactions -  
 PS Disclosure; SEQ ID NO 8868; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-ABL30511), expressed DNA  
 CC sequences (ABB57737-ABB72072),  
 CC (ABB57737-ABL16175) and the encoded proteins  
 CC  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX Sequence 593 AA;  
 SQ  
 Query Match 41.4%; Score 1389.5; DB 22; Length 593;  
 Best Local Similarity 49.0%; Pred. No. 1.9e-136;  
 Matches 284; Conservative 89; Mismatches 188; Indels 19; Gaps 10;  
 QY 39 IDTDLEAEPPEKRWVSNNIEFLMSCIATSVGLGNVWRFPFIAYONGGAFVLPYVYL 98  
 DB 3 LEESELSPDRKARDNMSSLEFLMSCIALSVGLGNVWRFPFIAYONGGAFVLPYVYL 62  
 QY 99 LVGGKPYVYLGCYVGFSSRSNVKWSISPMKGGYQAAGCCYIISYVYVTCGLCLYL 158  
 DB 63 VVGKPYVYEMLMGLQESSKGIQVDFAPLMRGVGAQLALGLATYVAVSVALTLRYF 122  
 QY 159 AMSFOATLPAWICQPEW-ENCVPSPDPTLASVNNITGTSSAQLYFRTVLOOSDGE 217  
 DB 123 FDSFASLELPWSPFCREMGDCGASGGQPLQGLSRNFSSYQIYLORIVLNETDSL 182  
 QY 218 LGAPWLVLCFTIAMLVGCVAVARGVSSGKAAYFALFPYVMTLFTTITILPGATD 277  
 DB 183 IGPPSGSLALMLGISWLTTLIIIRGVSSGKAAYVALFPYVMTLFTTITILPGATD 242  
 QY 278 GILFEVTPQAKLLELGVWTSVAVTQVFEFLTCTGPTIMSSVNGFPHNIRDAIVTTL 337  
 DB 243 GNVVFLTPQWKEKLEQVWYNATQVFFSLAVCFVITIMSSNRFGEHNYRDAIVTTL 302  
 QY 338 DTTSLSGCTTIGILGNATYELNS-EVGDVVGAGGTSIAFISYPAAIAK-TPQPOLFSV 395

DB 303 DTTSLSGCTTIGILGNATYELNS-EVGDVVGAGGTSIAFISYPAAIAKFKMPQVFSL 361  
 QY 396 LEFLMMSVYLIGIGSSVALLSTFNTLAMDAPFPRPVYMSAMTSCGFLGLVYCTPGGOYI 455  
 DB 362 LEFAMFLMGLVGSNVGMVSCIMTYLKDQFVAVKIMIIYVLSVIGFLVGLIYIRPGOHI 421  
 QY 456 LEIVDHGDTFLVLFCAISELAGVFWIYGLNCLDIEFMIGKKTGAVWRLCWGITPAI 515  
 DB 422 ITLMDFGVTFVSLVSAIFELIAGWIGYTRKLODAEYMINITSNYRCSIVTPVLY 481  
 QY 516 MTTFEYFALLASNNLVE-GDNY--VYPTAGY-VSGYLMFLFGMTFVPIGIGESLYKXRTG 571  
 DB 482 MLVILVYSLTMRPLSTNGQEFPLVYRVVGVCSGCT---IGOLF--YMAQIANFKOPKG 536  
 QY 572 TFSETTKAFHSKPSWGRSPRREREMWQFKAERAKALROK 611  
 DB 537 SLKSRINNSIKPHSDWGPDKPKIMDYQMF-----LRNK 570  
 RESULT 9  
 ID AAE05103  
 XX AAE05103 standard; Protein; 593 AA.  
 AC  
 XX AAE05103;  
 DT 18-SEP-2001 (first entry)  
 XX  
 DE Drosophila melanogaster dmsNF2.  
 XX  
 KM Fruit fly; invertebrate symporter cell surface receptor; dmsNF2;  
 XX sodium/neurotransmitter family; SNF; biopesticide; therapy.  
 XX Drosophila melanogaster.  
 XX  
 PN WO200149848-A2.  
 XX  
 PD 12-JUL-2001.  
 XX  
 PF 28-DEC-2000; 2000WO-US35551.  
 XX  
 PR 30-DEC-1999; 99US-0173929.  
 XX 15-MAR-2000; 2000US-0189399.  
 XX 23-MAR-2000; 2000US-0191686.  
 XX 23-MAR-2000; 2000US-0191687.  
 XX 23-MAR-2000; 2000US-0191688.  
 XX 23-MAR-2000; 2000US-0191695.  
 XX  
 PA (GENO-) GENOPTERA LLC.  
 XX  
 PI Kellerman KA, Keegan KP, Edens AJ, Torpey J;  
 XX  
 DR WPI: 2001-441879/47.  
 XX N-PSDB; AAD09683.  
 XX  
 PT Novel invertebrate symporter cell surface receptor proteins and nucleic  
 XX acid encoding the protein useful as pesticide or drug target and to  
 XX identify compounds that have utility as therapeutics or pesticides -  
 PS Claim 18; Page 65-66; 71pp; English.  
 XX  
 CC The invention relates to invertebrate symporter cell surface receptors  
 CC of the sodium/neurotransmitter family (SNF) and nucleic acid molecules  
 CC encoding such receptors. The SNF protein is useful for detecting a  
 CC candidate compound especially a putative pesticidal or pharmaceutical  
 CC agent that interacts with an invertebrate symporter cell surface  
 CC receptor protein or its fragment. Insect or worm genetically modified  
 CC to express protein of the invention is useful for studying invertebrate  
 CC symporter cell surface receptor protein activity, by detecting the  
 CC phenotype caused by the expression or mis-expression of the protein in  
 CC the animal. Nucleic acids encoding the invertebrate receptor protein or  
 CC their fragments are useful as biopesticides. SNF nucleic acids are  
 CC useful for generating mutant phenotypes in an animal model or living  
 CC cells that are used to study the regulation of genes encoding the







PF 20-AUG-1997; 97MO-US14637.  
 XX  
 PR 20-AUG-1996; 96US-0700013.  
 XX  
 PA (ALIX ) ALLELIX NEUROSCIENCE INC.  
 PI Albert VR, Borden LA, Kowalski LR, McKelvy JF;  
 XX  
 DR WPI: 1998-169158/15.  
 DR N-PSDB: AAV22906.  
 XX  
 PT Human glycine transporter - useful for treating nervous system  
 PC disorders, e.g. pain, myoclonus, etc.  
 XX  
 PS Disclosure; Pages 48-50; 90pp; English.  
 XX  
 CC The present sequence represents a variant of the human GlyT-2  
 CC transporter. The cDNA encoding GlyT-2 was generated by reverse  
 CC transcription PCR (RT-PCR) from human spinal cord or cerebellum polyA  
 CC mRNA. GlyT-2 is a glycine transporter found predominantly in the brain  
 CC stem and the spinal cord. GlyT-2 regulates the levels of glycine, a  
 CC major neurotransmitter, that preferentially influences the activity of  
 CC N-methyl-D-aspartate (NMDA) receptors and strychnine-sensitive glycine  
 CC receptors. The GlyT-2 protein can be used to treat or to diagnose a  
 CC nervous system disorder or condition, e.g. pain, myoclonus, muscle spasm  
 CC or hyperactivity, epilepsy or spasticity associated with stroke, head  
 CC trauma, neuronal cell death, multiple sclerosis, spinal cord injury,  
 CC dystonia, Huntington's disease or amyotrophic lateral sclerosis.  
 CC  
 SQ Sequence 797 AA;

Query Match 32.3%; Score 1084.5; DB 19; Length 797;  
 Best Local Similarity 38.2%; Pred. No. 3e-104;  
 Matches 221; Conservative 106; Mismatches 213; Indels 39; Gaps 10;

QY 54 MSNIEFLMCIATSVGLGNVWRPPFYAYONGGAFIPYVYVLLGKPPYVIECVLQ 113  
 DB 194 WSSKLDFILSMVGAAGLVGNWRPPYLAFOGGGAFILPYLMALAGLPFFLEVSIGQ 253  
 QY 114 FSSRNSYKWSISIPAMGTGYAQAAGCYILSYVVICGLVLYLAMSFOATLPMAICOP 173  
 DB 254 FASGGRPSYVKAIKPAIDLOGGCIAMILINSVLAIYINVTICTLFLFASFSVLPWGSNN 313  
 QY 174 EW-----ENCVPSD-PTL-----AASVNIT--NGTSSA-----QL 201  
 DB 314 PMWTPECKDKTKLLDSCVSDHPRKIQIKNSTFCMTAYPNVTWNTFSQANKTFVSGSEE 373  
 QY 202 YFLRTVYQSDGIEGGIGAPYVYVLCFLTAMLVFEGVAVRGVSSKAAFLALFPYVY 261  
 DB 374 YFKYEVKISAGIEYP-GEIRWPLALCLFLAMVIVAVSLAKGITSKVYVFTATFPYV 432  
 QY 262 MITLEFITTILPGATDGLFFVFPQMAKLELGVYSAVVOVPSFLVCTGPTIMFSYN 321  
 DB 433 LVILLINGVTLPGAGAGIWFITPKWEKLDATYMKAAQOIFESLSAANGILLTSSYN 492  
 QY 322 GFRHNIVRDAMVITTLDTFTSFGCTIFGILGNLAVELNSEVDVYAGGTSIAFTSY 381  
 DB 493 KFHNNCYRDLIYVCTNSATISIFAGFVIFSVIGEMANERKVINENVADQ-CPGIAFYVP 551  
 QY 382 DATAKTQPOLESVLFFLMSVIGIGSSVALLSTFNNLAMDAPFRVYVMSANT---CS 438  
 DB 552 EALRLPLSPFWALITFLMLTLGLDTPMATEIETIVISIDEPKRYTRHKPVFTLCCV 611  
 QY 439 CGEFLGLVYCTPGGOYLLEVDHGTGTFVLFCAISELAGVFIYGLNCLIDIEFMIG 498  
 DB 612 CFPIFMGPMTTGGIYVFOVLDYVYAAAYALVITAIFFLVGISTYVYGQRCEDLEMMIG 671  
 QY 499 KTGAYNRLCWGITPAITVTFEYVALLASNNLVFGDNVYVYVYAGVSGYTLMLFGMTF 558  
 DB 672 QPNFMKVCWAFVPTLITFLCFSSFYQWEPMTYV-SYRYPNMSMVLGIMLACSVIIMP 730  
 QY 559 IGIGFSIYKRYTGFSTIKKAFHSKPSNGPRSRERE 597

DB 731 IMFYIKMH-LAPGRFIERLKLVCSPQDPWGPFLAQHRGE 768

# RESULT 13

AAV82115  
 ID AAV82115 standard; Protein: 797 AA.

AC AAV82115;

XX 05-JUN-2000 (first entry)

DE Human SC polypeptide glycine transporter SEQ ID NO:2.

XX Human; glycine transporter; SC6 polypeptide; screening; diagnosis;

KW identification; neuropathic pain; spasticity; myoclonus; epilepsy;

KW stroke; head trauma; spinal cord injury; dystonia; multiple sclerosis;

KW amyotrophic lateral sclerosis; Huntington's disease;

OS Homo sapiens.

PN EP982400-A1.

PD 01-MAR-2000.

XX 27-OCT-1998; 98EP-0203615.

PR 28-AUG-1998; 98GB-0018890.

PA (SMIK ) SMITHKLINE BEECHAM PLC.

PI Brown AM, Chapman CG, Gloger IS, Evans JR, Cairns W, Heldon H;

XX WPI: 2000-258613/23.

DR N-PSDB: AA295536.

PT New isolated polypeptides and polynucleotides, useful for screening for  
 stroke, multiple sclerosis, Huntington's disease or Parkinson's disease

PS Claim 4; Page 17-20; 27pp; English.

XX The present sequence represents the human SC polypeptide glycine

CC transporter. The polypeptide is useful in a method for screening to

CC identify compounds which stimulate or inhibit the function of the

CC polypeptide and in a process for diagnosing a disease or a

CC susceptibility to a disease in a subject related to expression or

CC activity of the polypeptide e.g. treatment of neuropathic pain, other

CC pain, spasticity, myoclonus, epilepsy, stroke, head trauma, spinal cord

CC injury, dystonia, multiple sclerosis, amyotrophic lateral sclerosis,

CC Huntington's disease or Parkinson's disease.

XX Sequence 797 AA;

Query Match 32.3%; Score 1082.5; DB 21; Length 797;  
 Best Local Similarity 38.3%; Pred. No. 4.9e-104;  
 Matches 222; Conservative 105; Mismatches 213; Indels 39; Gaps 10;

QY 54 MSNIEFLMCIATSVGLGNVWRPPFYAYONGGAFIPYVYVLLGKPPYVIECVLQ 113  
 DB 194 WSSKLDFILSMVGAAGLVGNWRPPYLAFOGGGAFILPYLMALAGLPFFLEVSIGQ 253  
 QY 114 FSSRNSYKWSISIPAMGTGYAQAAGCYILSYVVICGLVLYLAMSFOATLPMAICOP 173  
 DB 254 FASGGRPSYVKAIKPAIDLOGGCIAMILINSVLAIYINVTICTLFLFASFSVLPWGSNN 313  
 QY 174 EW-----ENCVPSD-PTL-----AASVNIT--NGTSSA-----QL 201  
 DB 314 PMWTPECKDKTKLLDSCVSDHPRKIQIKNSTFCMTAYPNVTWNTFSQANKTFVSGSEE 373  
 QY 202 YFLRTVYQSDGIEGGIGAPYVYVLCFLTAMLVFEGVAVRGVSSKAAFLALFPYVY 261  
 DB 374 YFKYEVKISAGIEYP-GEIRWPLALCLFLAMVIVAVSLAKGITSKVYVFTATFPYV 432

```

QY      26  MILEFTTIIILPGATDGLIEFTPOMAKILBELGVYSAVQVFFSLVYCGPIITMESSYN 321
      :: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|
Db      433  LVLLIRGVLTIPGAGAGIYFTTPMEKLTJNATVWKDATTIEFSLSAAMGILLTLESYN 4920
      :: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|
QY      322  GFRHNITRDAMIYTTTLDFTTSPLESCITFGILIGNLAYENSEVDVYAGAGTSLAFSTYP 381
      :: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|
Db      493  KFINNCKYRDTLLVTCNSATSIFFACFVIFSVIYGEMANRKNIENTVADQ_GPGTIAFYVP 5511
      :: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|
QY      382  DALAKTFOPOLSVFEFLMASVYLGISSVALLSTFNTLAMDAPFRVPTVYMSANT---CS 438
      :: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|
Db      552  EALTRLPLSPFAIIFFMLTLTLDGTDEATLETIVTSIDSEPKYLRTHKPVPTLGGCI 6111
      :: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|
QY      439  CGELLGLVYCTGCGQYILEVDHRYGTLVFCALISELAGVFWIYGLENCLDIEFMIGK 498
      :: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|
Db      612  CFEIMEFPMITGGGIMFQYLDVTVAAASYALVIAIFELVGISSYYGQRFCEDEENMIGF 6711
      :: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|
QY      499  KTCGAYRLONGYITPAIMTTFVFAVLLASNNLVEDNDNVYVTAAGVSGYLMPLFGMTFVP 558
      :: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|
Db      672  QPPIEFKVCMAFATPTIILTFILCFSEFYQMEPTYG-SYRYPNWSMVLGMILMACSYTIP 730
      :: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|
QY      559  IGIGFSLVYKRTGTESETIKKAFHSKPSMGSPRPRERE 597
      :: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|
Db      731  IMEVIKMH-LAPGRFIERKLACASQDPWGPFLAHOHRE 768
      :: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|:: :|

```

RESULT 14  
AAB23323

AC AAB23323;

DT 23-FEB-2001 (first entry)  
VY

Human glycine transporter type 2 protein mutant #5.

KM Human glycine transporter type 2; GLYT2; muscle spasticity; myoclonus;  
KM epilepsy; stroke; head trauma; multiple sclerosis; spinal cord injury  
KM dystonia; Huntington's disease; amyotrophic lateral sclerosis;  
KM mutant; mutein.

OS	Homo sapiens.
OS	Synthetic.

PH	key	Location/Qualifiers
FT	Misc-difference	188
FT		/note= "wild-type Asn substituted by Asp"
yy		

PN WO200029564-A2

PD 25-MAY-2000.

PF 09-NOV-1999; 99WO-US26534.  
YY

PR 12-NOV-1998; 98US-0191468.  
YY

PA (GLIA-) GLIATECH INC.  
XY

PI Gallagher MJ, Burgess LH, Brunden KR; XY

DR WPI; 2000-387780/33.  
xy

PT New nucleic acid encoding a human glycine transporter type 2 (GLYT2),  
PT useful for expressing GLYT2 which can then be used for detecting  
PT whether an animal has autoimmune antibodies against a glycine  
PT transporter -

PS Claim 4; Page -; 175pp; English.

CC The present sequence is a mutant version of the protein sequence of the  
CC human glycine transporter type 2 (GLYT2). Glycine is involved in  
CC neurotransmission in the spinal cord and cerebellum, where glycine  
CC receptor inhibition results in pain transmission. The GLYT2 gene and

protein could, therefore, ~~not~~ be used to identify inhibitory compounds which prevent G1/2 activity, causing a decrease in muscle hypercontractility. They could be used to treat diseases or conditions associated with muscle contraction, including muscle spasticity (such as that due to epilepsy, stroke, head trauma, multiple sclerosis, neuronal cell death, spinal cord injury and dystonia), myoclonus, epilepsy, Huntington's disease, anisotropic lateral sclerosis and pain.

Note: This sequence is not shown in the specification, but is derived from the hGlyT2 wild type sequence shown in SEQ ID NO: 124 (AAMB09899).

SQ	Sequence	797 AA;
----	----------	---------

Query Match	32.2%	Score 1081.5	DB 21	Length 797
Best Local Similarity	38.1%	Pred. No. 6.2e-104		
Matches 223	Conservative 106	Mismatches 218	Indels 39	Gaps 10

QY 47 EPERWVWNNTEFLMSCIATSVGIGNWPRPEIYVONGGAFIYVYLLVKKPPY 106  
Db 187 EDARGNWSSKIDFLIISWGVAVGLNWRPEYFLAFONGGAFIYVYLLVKKPPY 246  
QY 107 LECVLGOFSSRSNWKVWISIPAMKGTGYAOAAGCYIISYYVIGLCLYYLAMSFOATL 166  
Db 247 LEVLSGQFASQGVSVWKAIPALOGGIGLMLIISLVLAIVYVILCIYTLFLFASFSVL 306  
QY 167 PAIACQPEW-----ENCYPSD-PPL-----AASVNNIT--NGTSSA-- 199  
Db 307 PMGSCNNPNWTECKDKRTKLLDSCISDHPKIQIKNSTFCMTAPVNTWNTFSOANKT 366  
QY 200 -----OLYFLRPLQOOSDIGEGGLAPMYIYLCIFIMLWFGVAVGWSCKRAAYL 25  
Db 367 FVSGSEIRKYFLYKISAGIEYR-GEIRPLALCFIEMVIVYASLAGITSGKYVEIT 422  
QY 255 ALFPYVMTLFTITTLIPGANDGILFEYTPQMAKLELGWYSAVTQVFFSLVTCGPI 314  
Db 426 AFPPYVYVILLIRGVTLPAGAGAMFYITPRMEXLTDLATYWKDAQNIQFISLAANGL 488  
QY 315 IMSSSINGRHRNIYRDAMVYTLUDLFTFSLSCTIFGLGNLAELNSEVDVYAGGTS 374  
Db 486 ITLSSYNNKHNHCYRDLTLVCTNSATSIFFAQFVFEISVIGMANERKNINENAVDO-GPG 544  
QY 375 LAEISYDAIAKTFQDFLSVLEFIMSMVSLIGGSSVALLSFENTLAMDAPRPVYMSA 434  
Db 545 IAAVYVPELTRLPLSPFAIITFFMLLTGLDITFAITETIIVISIDEPFKYRTHKPV 604  
QY 435 MT---CSCGELLGLVYCTPGGQYIILEVDHKGFTLVLECAISLSELAGFYVIGLENCLD 491  
Db 605 FTLGGCICFIFMGEPMTGGGGLYMQVDVYASALVIAITLPELVGSIYVYGQRCED 666  
QY 492 IEPMLGKTKGAWRLCQGVYTPAINTVFEYFALLASNNLVFGDNVYVPTAGVSGYIMLF 551  
Db 665 IEMMIGFOFNIMFKWCAWVPTIITLFIICSEFYQWEMPMYTG-SYRPNMSMVLGIMLMA 723  
QY 552 LGMTFVPIGIGSLKYRGTGSESEIINKAFHSPKSGRSPRERRE 597  
Db 724 CSYIMLPIMFVILKMH-LAPGRFETKULVCSBPQWPGFLEAONHGE 768

RESULT 15  
AAW56372

ID	AAW56372 standard; Protein; 797 AA
xx	

AC AAW56372

DT 29-JUL-1998 (first entry)  
xx

Amino acid sequence of a human GLYT-2 transporter protein

KM Human: *lyl-2* transporter; glycine transporter; neurotransmitter activity; N-methyl-D-aspartate receptor; NMDA; strychnine-sensitive glycine receptor; treatment; diagnosis; nervous system disorder; myoclonus; muscle spasm; hyperactivity; epilepsy; spasticity; head trauma; neuronal cell death; multiple sclerosis; spinal cord injury; Huntington's disease;

CC The present sequence is a mutant version of the protein sequence of the  
CC human glycine transporter type 2 (GLYT2). Glycine is involved in  
CC neurotransmission in the spinal cord and cerebellum, where glycine  
CC receptor inhibition results in pain transmission. The GLYT2 gene and

KW amyotrophic lateral sclerosis.  
 OS Homo sapiens.  
 PN W09807854-A1.  
 PD 26-FEB-1998.  
 PF 20-AUG-1997; 97WO-US14637.  
 XX 20-AUG-1996; 96US-0700013.  
 PA (ALX ) ALLELIX NEUROSCIENCE INC.  
 PI Albert VR, Borden LA, Kowalski LR, McKelvy JF;  
 DR WPI, 1998-169158/15.  
 XX N-PSDB; AAV22909.  
 PT Human glycine transporter - useful for treating nervous system  
 PT disorders, e.g. pain, myoclonus, etc.  
 PS Claim 1; Pages 55-56; 90pp; English.  
 CC The present sequence represents a human GlyT-2 transporter protein.  
 CC The cDNA encoding GlyT-2 was generated by reverse transcription  
 CC PCR (RT-PCR) from human spinal cord or cerebellum polyA mRNA. GlyT-2 is  
 CC a glycine transporter found predominantly in the brain stem and the  
 CC spinal cord. GlyT-2 regulates the levels of glycine, a major  
 CC neurotransmitter, that preferentially influences the activity of  
 CC N-methyl-D-aspartate (NMDA) receptors and strychnine-sensitive glycine  
 CC receptors. The GlyT-2 protein can be used to treat or to diagnose a  
 CC nervous system disorder or condition, e.g. pain, myoclonus, muscle spasm  
 CC or hyperactivity, epilepsy or spasticity associated with stroke, head  
 CC trauma, neuronal cell death, multiple sclerosis, spinal cord injury,  
 CC dystonia, Huntington's disease or amyotrophic lateral sclerosis.  
 XX  
 SQ Sequence 797 AA;

Query Match 32.2%; Score 1080.5; DB 19; Length 797;  
 Best Local Similarity 38.2%; Pred. No. 7.9e-104;  
 Matches 221; Conservative 106; Mismatches 213; Indels 39; Gaps 10;

QY 54 MSNNIEFLMSCIATSVGLGNVWRPFLAYONGGAFLLPYVLLVGRPYVLECYLQ 113  
 Db 194 WSKLDFILSNVGYAVGLGNVWRPFLAYONGGAFLLPYVLLVGRPYVLECYLQ 253  
 QY 114 FSSRNSYKWSISPAAGGTGAAAGGGLSYVYICGLIYLAISFOATLPAICOP 173  
 Db 254 FASGPPVSWKALPALGCGIAMIISVLAITYNVITCYLFLFASVSVLPWGCNN 313  
 QY 174 EM-----ENCVPD--PTL-----AASVNNIT--NGTSSA-----QL 201  
 Db 314 PMNTEPECKDKTKLLDSCVSDHDKIOIKNSTFCMTAYPNVTMYNFTSQANKTFVSGSEE 373  
 QY 202 YFLFTVLQOSDGLIEGGLAPIWYLVLCLFTIAMLVFGVAVGKSSGAAVFLALPPYV 261  
 Db 374 YFKFVLIKISAGIEXP-GEIRWPLALCLFLAMVIVYASLAKGKTSKVVYFFATFPYV 432  
 QY 262 MITLFTITILPGATDGLFFVTPQAKLLELGVYSAVTOVFSLVCTGPIIMSSYN 321  
 Db 433 LVILLINGVTLPGAGAIWYFTIPKWEKLDATYWKDAATOIFPFLSAAWGGTLTSSYN 492  
 QY 322 GFRHNIRDAIVITLDTFSLGCTIFGILGNLAVELNSEVGVAGAGTSLAFISYP 381  
 Db 493 KFRHNCTRDILITCTNSATISIFAGVIFSVIGFMANERKYNIEVADQ--GPGIAFVYIP 551  
 QY 382 DAIKTFQPOLFSVLFELMSVIGISSVALLSTFNTLADAPRPVTVYMSAMT---CS 438  
 Db 552 EALTRPLSPFMALIFLMLTLTGIDMPATIFETIVTSISDEPKYLRTKHVPFTLGGCI 611  
 QY 439 CGELGLVYCTPGGQYILEVDHKGTFVLFCALISLAGVFWTYGLENLCIDTEPMGK 498

Db 612 CFFIMGFPMTGGIYWFQLVDTYAASTALVIAIFELVGISVYGLQRECEDEIMMIGF 671  
 QY 499 KTGAVRMLCQGVTPAINTVFEFFYALLASNNLVGDNVYVPTAGYSGYMLFLGMTEVP 558  
 Db 672 QPNIEMKCAVYPTILFIFLCSFYQWEPMTYG-SYRIPNMSAVLGLMLACSVIWP 730  
 QY 559 IGIGFSLYKRYGTGFSETIRKAFHSKPSWGRPSRPRERE 597  
 Db 731 IMFVIMKH-LAPGRFIERLKVCSQPDMDGPFLLAQHRRGE 768

Search completed: July 21, 2003, 09:29:29  
 Job time : 74 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2003, 09:24:45 ; Search time 43 Seconds  
(without alignments)  
1415.188 Million cell updates/sec

Title: US-09-991-458-2

Perfect score: 3354

Sequence: 1 MNDQVNGGESESEPKMEPK.....TSRVKHLWYSITGAVRRNIN 633

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1076.5	32.1	799	2 A48716	glycine transporte
2	1027	30.6	637	2 JH0674	L-proline transporte
3	1011	30.1	599	2 S11073	gamma-aminobutyric
4	1008	30.1	599	1 ACRGT	gamma-aminobutyric
5	999	29.8	598	2 A51368	gamma-aminobutyric
6	991	29.5	630	2 A47398	serotonin transport
7	981	29.2	630	2 S30604	neurotransmitter t
8	979	29.2	611	2 S46260	creatine transport
9	978.5	29.2	627	2 P46027	gamma-aminobutyric
10	973	29.0	627	2 JH0695	gamma-aminobutyric
11	972	29.0	635	2 S23431	choline transport
12	971.5	29.0	607	2 S19585	serotonin transport
13	971.5	29.0	635	2 G02095	creatine transport
14	971	29.0	635	2 J02386	creatine transport
15	970.5	28.9	635	2 A46061	Na(+)-dependent cr
16	968.5	28.9	620	2 S46487	taurine transporte
17	968.5	28.9	620	2 G01426	taurine transporte
18	967	28.8	692	2 I57956	glycine transporte
19	966	28.8	655	2 A46270	Na and Cl dependen
20	965	28.8	602	2 A45078	gamma-aminobutyric
21	964	28.7	614	2 A43390	gamma-aminobutyric
22	963	28.7	627	2 B44409	gamma-aminobutyric
23	959	28.6	602	2 A44409	gamma-aminobutyric
24	956	28.5	638	2 I77912	glycine transporte
25	954.5	28.5	619	2 S29839	taurine transport
26	954.5	28.5	622	2 S63539	GABA/beta-alanine
27	953.5	28.4	621	2 I57939	taurine transporte
28	949.5	28.3	620	2 A48980	dopamine transport
29	946.5	28.2	620	2 I57937	dopamine transport

30	940.5	28.0	614	2 A41757	betaine transport
31	936	27.9	615	2 S43285	norepinephrine tran
32	935	27.9	633	2 I58140	glycine transporte
33	935	27.9	633	2 S23151	glycine transporte
34	935	27.9	638	2 JH0673	glycine transport
35	934.5	27.9	619	2 S20346	dopamine transport
36	933	27.8	619	2 I59558	dopamine transport
37	932.5	27.8	614	2 S68236	betaine/GABA trans
38	929.5	27.7	590	2 A47194	taurine and beta-a
39	929.5	27.7	597	2 S65673	GABA transport pro
40	929	27.7	602	2 I55651	norepinephrine tran
41	920.5	27.4	617	2 S14278	norepinephrine tran
42	902	26.9	693	2 A41617	dopamine transport
43	867.5	25.9	579	2 D88551	protein T23G5.5 [1
44	861	25.7	615	2 T43330	catecholamine tran
45	820	24.4	632	2 T24405	hypothetical prote

## ALIGNMENTS

## RESULT 1

A48716  
glycine transporter GLYT2.- rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Jun-1995 #sequence\_revision 02-Jun-1995 #text\_change 02-Jun-1995  
C:Accession: A48716  
R:Lin. Q.R.; Lopez-Corcuera, B.; Mandiyan, S.; Nelson, H.; Nelson, N.  
J. Biol. Chem. 268, 22802-22808, 1993  
A:Title: Cloning and expression of a spinal cord- and brain-specific glycine transport  
A:Reference number: A48716; MUID:94043045; PMID:8226790  
A:Accession: A48716  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-799 <LID>  
A:Cross-References: GB:I21672  
A:Genetics:  
C:Gene: GLYT2  
C:Keywords: brain; transmembrane protein

Query Match	32.1%	Score 1076.5;	DB 2;	Length 799;
Best Local Similarity	37.5%	Pred. No. 1.1e-74;		
Matches	224;	Conservative 107;	Mismatches 228;	Indels 39;
Gaps	10;			
QY	35	ALDNIDDDLEAPERRVNSNNIEFMSCTSVGLGVWMPFPPIAYONGGATLPYV	94	
DB	177	ATTTPEDQGDENKARGNWSKLPDITSMVAVGLGVWMPFPPLAFQNGGATLPYL	236	
QY	95	IYLLVGRVYVLECEVGFSSRNSGVKWSISPAKMGTYAQAACGYLSYVYICGL	154	
DB	237	MMALALAGLPFFLEVLSTGFPASQGPVSWKATPALQGCIALMLIISVLAITYNNICT	296	
QY	135	LYLLMSFOATLPMALICQPEW-----ECVPSD-PTL-----AASVNI	192	
DB	297	LEFLASFVSVLPMGSCNNPMTPECKDKTKLLSCVIGDHPKIQIKNSTPCMAYPNL	356	
QY	193	T--NGTSSA-----OLYELRTVLOOSDGEIGLGAPIWVLCIFPAMLVGVYAR	242	
DB	357	TIWNTTSQANKTFVSGSEYFVKYFVAKISAGIEYR-GETRWLPCLFLAMIVYASLAK	415	
QY	243	GVKSSGKAAYFLALPEYVVMITLFTTIIIPGATDGIIEFVTPOMAKLLELVMSAVTQ	302	
DB	416	GKSGKGVYFATPEYVVLVLLRGVLPAGAGIWEFIPKMEKLDATVMDAATQ	475	
QY	303	VFSITVTCGPIIMSSYVGFHNNIYRQAIYTTDTFSLSCCTIRSIGLNLAYELNS	362	
DB	476	IFFSLSAMGGLITLSSYVKNFNNCYRDLIYTCNSATSIAGVIVSVIGFMANERV	535	
QY	363	EVGDVVGAGGTSLAFISYDPAIAKTFQQLFVFLFIMMSYGISSVALSTFPTLAMD	422	
DB	536	NLENVADQ-GPGIAVYVPEALTRPLSPFMAIIFFLMLITGLDMRTATITTYTSISD	594	
QY	423	APRPVPTVYMSAMT--CSCGFLIGLVYTCGGOYILLEVHVGSTFVLFCALSELAGV	479	

Db 595 EEPKRLTRKPEVFLGCCICFFIFGPMITGGIYMQLVDYTAASALVITALELVI 654  
QY 480 FWIAGLNLCDIEFMKKGKAGVRLCWMGITPAIMTVTFYFLLASNNLVEGDNYVP 539  
Db 655 SYVGLDRFCDDIMMIGFOPNITKWCMAVPTLITLFCFSFYQWEPMTYG-SYRYP 713  
QY 540 TAGVSGYLMFLMGTVPITGIGSLYKRYTGFSEITIKAFHSPKSPGPRSPRRE 597  
Db 714 NMSVVLGMLMAGSVIMPIFVTKM-LAGRIEHLKLVCSQPDMGPFLLAQRGE 770

## RESULT 2

JH0674  
L-proline transport protein - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 28-Feb-1997  
C:Accession: JH0674  
R:Freeman Jr., R.T.; Caron, M.G.; Blakely, R.D.  
Neuron 8, 915-926, 1992  
A:Title: Molecular cloning and expression of a high affinity L-proline transporter expe  
A:Reference number: JH0674; MUID:92265305; PMID:1350201  
A:Accession: JH0674  
A:Molecule type: mRNA  
A:Residues: 1-637 <FR>  
A:Experimental source: forebrain  
C:Superfamily: gamma-aminobutyric acid transporter  
C:Keywords: glycoprotein; leucine zipper; phosphoprotein; proline transport; transmembr  
F:46-65/Domain: transmembrane #status predicted <TM1>  
F:73-93/Domain: transmembrane #status predicted <TM2>  
F:76-97/Region: leucine zipper motif  
F:117-137/Domain: transmembrane #status predicted <TM3>  
F:217-236/Domain: transmembrane #status predicted <TM4>  
F:242-264/Domain: transmembrane #status predicted <TM5>  
F:291-311/Domain: transmembrane #status predicted <TM6>  
F:323-345/Domain: transmembrane #status predicted <TM7>  
F:374-397/Domain: transmembrane #status predicted <TM8>  
F:424-443/Domain: transmembrane #status predicted <TM9>  
F:457-479/Domain: transmembrane #status predicted <TM10>  
F:500-519/Domain: transmembrane #status predicted <TM11>  
F:537-559/Domain: transmembrane #status predicted <TM12>  
F:1/Binding site: phosphate (thr) (covalent) (by CAMP-dependent kinase) #status predict  
F:41/Binding site: phosphate (thr) (covalent) (by protein kinase C) #status predicted  
F:132/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:239/600/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predict

Query Match 30.6%; Score 1027; DB 2; Length 637;  
Best Local Similarity 39.8%; Pred. No. 5,3e-71;  
Matches 227; Conservative 90; Mismatches 236; Indels 18; Gaps 11;

QY 41 DTDLAE-PPERWVSNINIEFLMSCIATSVGLNVRFPFLAYONGGAFLVPYIVLL 99  
Db 26 DVIDDDVFAADRGWMTKLDLSCICVGLGNVWFRPYRAYNGGAFLLVPEFLMAI 85  
QY 100 VGRPVYLLCEVLGOFSSRNYSKSTSPAKKGTGYAQAAGCYLLSYVYICGLTYLA 159  
Db 86 CGIPLEFLSLGQFSSGLPAAWKISPLERGAALLLVGLVAITYNMILAVLYLP 145  
QY 160 MSEFATLPMAICPEW--ENCV---PSPDLAASVNNITNGSSAOLYFLRYL--QOS 211  
Db 146 ASLTSNLPMEHCGMMWTERCLEHRPKDNGALPL-LSSTYSPSEVSRVLAHQGS 204  
QY 212 DGBGLGADPITVLCFLTAMLVGVAVGKSSKAAVFLALFPYVMTLFTTTII 271  
Db 205 QGI-GRGELIRNMLCTCLLAWIVFLCIIKGVSSSGKVVYFATPFYLLMLLVGYT 263  
QY 272 LPGAQGLFFVTPQMAKLELGWYSAYVQVFESLVCGPITMESYNGFRNITRDA 331  
Db 264 LPGAQGLFFVTPQMAKLELGWYSAYVQVFESLVCGPITMESYNGFRNITRDA 331  
QY 332 WITVTLDTFSLSGTIFGLILNLAVELNSEVDVGVAGGTSLAFISYDATAKTPQ 391  
Db 324 FYITLGAITSILAGFAIFSVLGMSELGVV-DQVAKAGPGLAFIYIQAAMTMLPSP 382

QY 392 LFSVLEFLMSVLGIGSSVALLSTENTLAMDAP---RVPTVWSAMTSCGFLGLVYC 448  
Db 383 FWSLEFLFMTLLTGIDSOFALETITVAVDEFPYLPKPAVNSGLCVAMITMLLT 442  
QY 449 TPGQVITELVDHVGTFVLFCAISBELAGVWITGLENCLDIEFMKKGKAGVRLCW 508  
Db 443 TDGGMVNLVLLDDVSASFGLWVVTTCIAVTRVYGIQRFCRDIMHMLGFKPGLYFRACW 502  
QY 509 GVTPAINTVTFPALLASNLVFGDNYVTFAGYSGYLMFLGMTFVPIGIGFSLYK 568  
Db 503 LFLSPATLLALLVYSIVKYQPEXG-SYRFPAMAEILGLMLGLSLMIPAGMLVAVLR- 560  
QY 569 RTGFSEITIKAFHSPKSPGPRSPRREWM 599  
Db 561 EGGSLMERLQOASRPALDMGP-SLEENRTGM 590

## RESULT 3

S11073  
gamma-aminobutyric acid transport protein - human  
N:Alternate names: GABA transporter  
C:Species: Homo sapiens (man)  
C>Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Aug-1999  
C:Accession: S11073  
R:Nelson, H.; Mandiyan, S.; Nelson, N.  
FEBS Lett. 269, 181-184, 1990  
A:Title: Cloning of the human brain GABA transporter.  
A:Reference number: S11073; MUID:90335367; PMID:2387399  
A:Accession: S11073  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-599 <NE>  
A:Cross-references: EMBL:X54673; NID:931657; PIDN:CA83484.1; PID:931658  
C:Superfamily: gamma-aminobutyric acid transporter  
C:Keywords: transmembrane protein

Query Match 30.1%; Score 1011; DB 2; Length 599;  
Best Local Similarity 35.4%; Pred. No. 8,3e-70;  
Matches 213; Conservative 118; Mismatches 229; Indels 42; Gaps 13;

QY 3 DQVNGFSESEPEMERKSSQIS-LPPANKALNDITDLEAEPPERWVSNINIEFL 61  
Db 10 DQGI-----STEVSEAVANDKPTLVYKQKKAADLPDRDTWGRDFL 54  
QY 62 MSCIATSVGLGNVRFPFLAYONGGAFLVPYIVLLVGRPVYLLCEVLGOFSSRNYSK 121  
Db 55 MSCVGYAIGLGNVRFPFLAYONGGAFLVPYIVLLVGRPVYLLCEVLGOFSSRNYSK 114  
QY 122 VMSISPAKGTGYAQAAGCYLLSYVYICGLTYLAMSFQATLPMAICPEW--ENCV 179  
Db 115 VMLKAPMFKGVLAAALSLPMLNITYIIVISMAIYLYNSPTTLPMKQCDNPMTDRCF 174  
QY 180 PSDPDLAASVNNITNGSSAOLYFLRTVLDQSDIEBGLAPITVYVLCFLIAMLMPGV 239  
Db 175 SN-----YSWNTNMTNSAVAEFERNMHMDTDLKP-QDIRPLATITTAIAWITVFC 228  
QY 240 VARGVSSGAAYVFLALFPYVMTLFTITLIEGADGLFEVTPQMAKLELGWYSA 299  
Db 229 IMKVGMTGKAVYTSATPIIMLITLIFRGVTLPGAEGILFTITPFRKLSDBEVLLDA 288  
QY 300 VTQVFSLVCTGPIIFESSYNGFRHNIYRDAMVITLDTFTSLGCTIFGLGNLAYE 359  
Db 289 ATOFFSYGLGLSLALAGSYNSPHNNVYRDSITVICINCSMTMFAFVFSIYGFMAHV 348  
QY 360 LNSVGVNAGAGSLAFISYPOAIATPQOLFVFLPMLMSVLGIGSSVALLSTFNTL 419  
Db 349 TKRSIAD-VAASGGLAFLAVPEAVYQLPISPMALLIFSMMLMLTGIDSOFCVTEGRTA 407  
QY 420 AMDAFPRV---PVTYVSAMTSCGFLGLVYCPTGGQVITELVDHVGTFLL-VLFCFAS 474  
Db 408 LVDEYPLLNRRRELPIAAV-CITSYLIGLSNITGGIYVFKFLFDIYSASGMSLLPLVFF 466



Db 347 VTNRPAD-VAASGPGALFLAPYATQPLSPMSLTFMSMLMLDIDSOFCVBEFIT 405  
Qy 419 LAMDAFRV---PTVMSAMTSCGFLGLVYCTPGQYVILEVDHYGFTL-VLECAI 473  
Db 406 ALVDEFKLLGRREIFL-AMVCIVSYGLISNTQGLYKFLDFDYSSAGMELLEVE 464  
Qy 474 SELAGVWYIGLEMLCDIEFMLGKTKGAYWRLCMGVITPAIMTVFEYALASNNLVFG 533  
Db 465 FEYVISMCGVNVNFENFNIEMVGHKPCMLMKLCHNSFTPLIVGVGLFSAIQMKPLMG 524  
Qy 534 DNYVYPAGYVGYLMLFLGNTFVPDIFGFSLYKRT-GTSEETIK 578  
Db 525 -SYTFPKMGQGVGMFALSSMLIP---GYMGVFLTSKSLKQRLR 567

RESULT 6  
A47398  
serotonin transporter - human  
N:Alternate names: 5-hydroxytryptamine transporter  
C:Species: Homo sapiens (man)  
C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Aug-1999  
C:Accession: A47398; S37688  
R:Ramamoorthy, S.; Bauman, A.L.; Moore, K.R.; Han, H.; Yang-Feng, T.; Chang, A.S.; Ganap  
Proc. Natl. Acad. Sci. U.S.A. 90, 2542-2546, 1993  
A:Title: Antidepressant- and cocaine-sensitive human serotonin transporter: molecular cl  
A:Reference number: A47398; MUID:93211998; PMID:7681602  
A:Accession: A47398  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-630 <RAM>  
A:Cross-references: GB:L05568; NID:q291819; PIDN:AAA3492.1; PID:g291820  
A:Experimental source: Placental trophoblastic cell line JAR  
A:Note: sequence extracted from NCBI backbone (NCBI:128282, NCBI:128283)  
J. Leach, K.P.; Molozin, B.L.; Estler, H.C.; Murphy, D.L.; Riederer, P.  
R. Neurol. Transm. 91, 67-72, 1993  
A:Title: Isolation of a cDNA encoding the human brain serotonin transporter.  
A:Reference number: S37688  
A:Accession: S37688  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-630 <LESS>  
A:Cross-references: EMBL:X70697; NID:q36432; PIDN:CAA50029.1; PID:g36433  
A:Note: the authors translated the codon GGA for residue 25 as Val and GAA for residue 5  
C:Genetics:  
A:Gene: GDB:SLC6A4; HTT  
A:Cross-references: GDB:134713; OMIM:182138  
A:Map position: 17q11.1-17q12  
C:Superfamily: gamma-aminobutyric acid transporter  
C:Keywords: transmembrane protein

Query Match 29.5%; Score 991; DB 2; Length 630;  
Best Local Similarity 37.0%; Pred. No. 3e-68;  
Matches 226; Conservative 97; Mismatches 242; Indels 46; Gaps 15;

Qy 1 MNDQVNGGSESEPEKPKRKSQISLPPANNKAALDNDIDTDLAE--PPERWWSNNT 58  
Db 38 VESGQISNGY-SAVPSGAGDDTRHSIPAT-----TLVAELHGERETWCKV 86  
Qy 59 EFLMSCATGAVGNGWRPEFLAYONGGAFVLYVYVLLVNGRPVYLLCEVLCGQFSRN 118  
Db 87 DELLSYIGTAVDGNWRRPYICYQNGGAFLLPYTMALEFGSIPLEMYELAGQYHRNG 146  
Qy 119 SVKVV-SISPMKGTGYAQAAGCGYLLSYVYVIGCLVLYLAMSFOATLPWALQPEWE- 176  
Db 147 CISIMWKICIFIGIGALICIIAFYASYNTIMANALYLLISSFDQLEPMTSCKNMNT 206  
Qy 177 -KCVSPDPLTASVNNIT---NGTSSAQLYFLTVLQ--OSDIEGGLGAPVYVLCIF 230  
Db 207 GNC-----TNFSESDITWTLSHSTSPAEFYTRHYLQIHRSKQLQ-DLGISWQALCIM 260  
Qy 231 IALWLVGVYARGVSSGKAAYFLALFPVYVMTLFTITITILGARDIGLIFETTPQAKL 290

Db 261 LITVYIESTIKKVKYTGKVVWVATPPIIILSVLYRGATLPBAMRGVLYLPKNOKL 320  
Qy 291 LELGVWYSAVTVQFSLTCTGTPIIMSSYNGFRHNTYRDAMIYTLDTFTSLSGCTIF 350  
Db 321 LETGVWIDAAQIFESLGPFGVYLAASVYKFNENNCYQDALVTSVNCMTSEFVSQVIF 380  
Qy 351 GIIGNLAYELNSEGVDVAGGTSIAFISPDIAKTFQQLFVLEFFIMVSGIGSSV 410  
Db 381 TVLGAMEMRREDVSEVAKDAGSLLFTYAEALANMPASTFAITFLMLITGLDSTIF 440  
Qy 411 ALLSTFNTLAMDAPRV-----PTVMSAMTSCGFLGLVYCTPGQYVILEVDHYG 464  
Db 441 AGLEGVITAVLDEEPHWAKRERFVAVITCFEGLVLTLE--CGAVVKKLEBYAT 497  
Qy 465 TFLVFCALSELAVFYIGLEMLCDIEFMLGKTKGAYWRLCMGVITPAIMTVFEYAL 524  
Db 498 GPVALTVALLAVAVSMFYGITQPCRDVKEMLGFSPEWFKICVWASPLFLFTICSFL 557  
Qy 525 LASNNL-VFGDNYVYPAGYVGYLMLFLGNTFVPDIFGFSLYKRTGTSEETIKAFHS 583  
Db 558 MSPQLRLF--QVNPYMSIILGICITSSFCIPYIAVRLT-TTGTRERLIK---- 610  
Qy 584 KPSWGPSPRE 594  
Db 611 -SITPETPE 619

RESULT 7  
S30604  
neurotransmitter transport protein - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Aug-1999  
C:Accession: S30604  
R:Mayser, W.; Beyer, H.; Schloss, P.  
FEBS Lett. 295, 203-206, 1991  
A:Title: Isolation of cDNAs encoding a novel member of the neurotransmitter transport  
A:Reference number: S30604; MUID:92111740; PMID:1765155  
A:Accession: S30604  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-630 <NAV>  
A:Cross-references: EMBL:X63995; NID:q56779; PIDN:CAA45401.1; PID:g56780  
A:Superfamily: gamma-aminobutyric acid transporter  
C:Keywords: transmembrane protein

Query Match 29.2%; Score 981; DB 2; Length 630;  
Best Local Similarity 37.2%; Pred. No. 1.8e-67;  
Matches 226; Conservative 94; Mismatches 241; Indels 46; Gaps 15;

Qy 5 QVNGGSESEPEKPKRKSQISLPPANNKAALDNDIDTDLAE--PPERWWSNNTIEFLM 62  
Db 42 QISNGY-SAVPSYASGEASHISIPAT-----TLVAELRGERETWCKMDFFL 90  
Qy 63 SCATISVGLGNWRPEFLAYONGGAFVLYVYVLLVNGRPVYLLCEVLCGQFSRNSVKY 122  
Db 91 SVIGYAVDIGNIRFPYICYQNGGAFLLPYTMALEFGSIPLEMYELAGQYHRNGCISI 150  
Qy 123 W-SISPMKGTGYAQAAGCGYLLSYVYVIGCLVLYLAMSFOATLPWALQPEWE--NCV 179  
Db 151 WKRICTPFKIGTICIAIYASYNTITAMALYLLISLTDLRLPMTSCNMGNC- 209  
Qy 180 PSDPTLAASVNNIT---NGTSSAQLYFLTVLQ--OSDIEGGLGAPVYVLCIFLAWL 234  
Db 210 ---TNFADNDITWTLSHSTSPAEFLRVHVLQIHSKGLQ-DLGTSIMQLTLCIYLIF 264  
Qy 235 MIVGVYARGVSSGKAAYFLALFPVYVMTLFTITILPGATGILFEVTPQAKLELG 294  
Db 265 VIFESIMKGVKTSKVVWVATPPIYVLSVLVKGATLPBAMRGVYVLYLPKNOKLELG 324  
Qy 295 VVYSATVQVFEVSLTCTGTPIIMSSYNGFRHNTYRDAMIYVTLDTFTSLSGCTIFGIG 354  
Db 325 VWYDAAQIFESLGPFGVYLAASVYKFNENNCYQDALVTSVNCMTSEFVSQVIFTVLG 384



QY 355 NLAELNSENVDVVGAGTSLAFISYPAIAKTFOPOLFSVLEFLMAGVIGISSVALLS 414  
DB 385 YAAEMENEDVSEVAKADGSLFFITAEAIAMNPSTFFAIIFFLLITLIGLSTAGLE 444  
QY 415 TINTLAMDAPPRVPT-----YMSAMTSCGFLGLVYCTPGGQYILEVDHYGGTFLV 468  
DB 445 GVTITAVLDFPHIMAKRREMFVLIIVITCVLSLLTL---TSGAVAVVTLLEBYANGP 501  
QY 469 LFCALISELAGVMIYGLNLCIDIEFMLGKKTGAWRICMGVITPAIMTTFEYALASN 528  
DB 502 LVALLEAVANAVMFIGITFCSDVKEMLGFSPGMWRICWVAISPLFLITCSFLMSP 561  
QY 529 NT-VFGDNVYPTAGVSGYMLFLGMFVPIGIFSLYKRYRGTSETIKRAFHSKPSM 587  
DB 562 QLRLEF--QYNYPHMSIVLGCIGMSSVICIPYIIYRLIS-TPGTLKEKIR-----SI 612  
QY 588 GPRSPRE 594  
DB 613 TPEPTPE 619

## RESULT 8

S46260

creative transporter - marbled electric ray

C:Species: Torpedo marmorata (marbled electric ray)

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Aug-1999

C:Accession: S46260

R:Gulabal, C.; Kilmann, M.W.

J. Mol. Biol. 241, 317-324, 1994

A:Title: A creative transporter cDNA from Torpedo illustrates structure/function relation

A:Reference number: S46260; MUID:94334996; PMID:8057375

A:Accession: S46260

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-611 &lt;G&gt;

A:Cross-references: GB:X77306; NID:9541753; PIDN:CA54512.1; PID:9541754

C:Superfamily: gamma-aminobutyric acid transporter

## Query Match

29.2%; Score 979; DB 2; Length 611;  
Best Local Similarity 35.1%; Pred. No. 2.4e-67;

Matches 208; Conservative 103; Mismatches 235; Indels 46; Gaps 10;

QY 47 EPEPMWNSNIEFLMSTCLASVIGLGNWRPFPIAYONGGAFIVPYVILLVGRPYV 106  
DB 32 EPGSVHTWSRQDFIMSCVFAVGLGNWRPFYLCYKNGGVFLPIYLLVAVFGDIPFE 91  
QY 107 LECVIGPSSRSNVWMSISPAKMGTYAQAAGCYIISYYVIGLGLIYLAISFOATL 166  
DB 92 LEISIGPMAKAGIINAMNIAFLFKGLGYASVYFFCNITVYLLVWSSFFLYVQSFFSL 151  
QY 167 FWAICQPEMENCVPSPDPTLAASVNNITNGTSSAQ----LYFLKTVLQOSDGIIEGIGAPI 222  
DB 152 FWASCNNTWNTAACYEAGANAS-TEIYPTAPAGSIYVOFWERAVLRLLSSGL-GVYGEIG 209  
QY 223 WTVLYICLFIAMLVGVAVARGVSKGAAYFLALFPYVMTLFTTTIILPGATDGIILF 282  
DB 210 WELTICLAIYMLVFCIMKGVKTSKAVYVATFPYIILVILLKRGVTLRGAOYGIYY 269  
QY 283 VTPQWAKLLEGVWTSAYTQVFFSLVCTGPILMSSYNGFRHNIRDAWITVTLDTPTS 342  
DB 270 LQPMGKLGEMQWVIDAQIIFFSYALIGLITLALGSLNOLHNDCKRAFIILSVNSATS 329  
QY 343 FLSCGTIGILGNLAELNSENVDVVGAGTSLAFISYPAIAKTFOPOLFSVLEFLMMS 402  
DB 330 FFAGLVVPSIILGFMAVEGVDI-SVAESGGLAEIAYPKAVITLMPFQVAVLFEIMLL 388  
QY 403 VLGIGSSVALLSTFNTLMDAPF-----RVPYVMSAMTSCGFLGLVYCTPGGQYILE 457  
DB 389 CLGIGSOVGVYEGVITAILDMPKFSFRYLRVYVAVAVICISFLIDISMTTEGMYIFQ 448  
QY 458 LVYH--GGTFLVFCALISELAGVFIYGLNLCIDIEFMLGKKTGAWRICMGVITPAI 515  
DB 449 IFDYYSASGTTL-LMTAFMECVAVAVMYGGRYIDDLAMNLGYRPMALVKKWCSYITPLV 507

QY 516 MTTVFYFALLASNNLVFGDNVYPTAGVSGYLMFLGTEVPIGIFSLYRGTSE 575  
DB 508 CMGIFTFEHLVYKPLVNTYTYTPWVGAEIGCLALASMLCVPTVYLSLSGR-GSLKE 566  
QY 576 TIKKAFHSKPSMGPSRFREREMQFKAKAKLRQKMNSRYKHLWYSITGA 627  
DB 567 RWRKRL--TTPVWASH-----HLAYKWA 588

## RESULT 9

F46027

gamma-aminobutyric acid transporter - mouse

C:Species: Mus musculus (house mouse)

C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 16-Feb-1997

C:Accession: F46027

R:Liou, O.R.; Mandiyan, S.; Nelson, H.; Nelson, N.

Proc. Natl. Acad. Sci. U.S.A. 89, 6639-6643, 1992

A:Title: A family of genes encoding neurotransmitter transporters.

A:Reference number: A46027; MUID:92335351; PMID:1631167

A:Accession: F46027

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-598 &lt;LID&gt;

A:Experimental source: brain

C:Superfamily: gamma-aminobutyric acid transporter

## Query Match

29.2%; Score 978.5; DB 2; Length 598;  
Best Local Similarity 34.3%; Pred. No. 2.6e-67;

Matches 208; Conservative 123; Mismatches 232; Indels 43; Gaps 13;

QY 3 DGQVNGGESSEPKMKPRSSQIS-LPPANAKALDNIDDDLEAPEPMWNSNIEFL 61  
DB 10 DGQI-----STEVSEAPVASDKPKTYLVKVKAKADLPDRDWRKGRPFL 54  
QY 62 MSCIAVGLGNWMPFPIAYONGGAFIVPYVILLVGRPYVILECVLOFSSRSNVK 121  
DB 55 MSCVYAIAGLGNWMPFPIAYONGGAFIVPYVILLVGRPYVILECVLOFSSRSNVK 114  
QY 122 VMSISPAKMGTYAQAAGCYIISYYVIGLGLIYLAISFOATLPMACIPEW--ENCY 179  
DB 115 VNNMAMPKRGVAVAAVLSFPMNIIYIYIISAIYILNLSFTTILPMKQCONPMNTDCE 174  
QY 160 PSDPTLAASVNNITNGTSSAQIIFLRTVLOOSDGIIEGIGAPIWTVLYICLFIAMLV 239  
DB 175 SN-----YSLVNTTMTSAVVEFWERNNHOMTDGID--KPGQIRCLAITLIAIAYLYFC 227  
QY 240 VARGVSSGKAAYFLALFPYVMTLFTTTIILPGATDGIILFPVPMOKAKLLEGVMSA 299  
DB 228 IMKGVGMKGVYFSAITPYIMLILFLFRGVTLPGAKGILFTYIPNRKLSDSSEVTPDA 287  
QY 300 VTQVFFSLVCTGPILMSSYNGFRHNIRDAWITVTLDTPTSLSGCTIGILGNLAVE 359  
DB 288 ATQIFFSYGLGIGSLAISGSYFNHNYRDSIIYCCINCSMTAGVIFSIYGFMAHV 347  
QY 360 LNSEVDVVGAGTSLAFISYPAIAKTFOPOLFSVLEFLMAGVIGISSVALLSTFNTL 419  
DB 348 TKRSIAD-VAASGGLAEIAYPKAVITLMPFQVAVLFEIMLLMIGIDSDCTVEGFTTA 406  
QY 420 AMDAPPRV-----PYVMSAMTSCGFLGLVYCTPGGQYILEVDHYGGTFL-VLFCATS 474  
DB 407 LVDEYPRLLRRRRELFIAAV-CIYSYLLGSLNITNOGGIYVKKLDYYSASMSLLFLVFF 465  
QY 475 ELAGVFIYGLNLCIDIEFMLGKKTGAWRICMGVITPAIMTTFEYFALLASNNLVFGD 534  
DB 466 ECVSISMPYGVNREYDNIQEMVSGSPCIWMKLCWSPFPPIIVAGVFLFSAVQMPPLMG- 524  
QY 535 NYVYPTAGVSGYLMFLGTEVPIGIFSLYKRYRGTSETIKRAFHS-----RPSW 587  
DB 525 SYVPPKMGQGVIMLALSSVLLP---GIMAYMFLTLKGLSKLQKRVNIOQSDIYVREN 581

Oy 588 GPRSPR 593  
Db 582 GPEPQ 587

## RESULT 10

gamma-aminobutyric acid transporter protein 3 - rat  
JH0695  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 20-Aug-1999  
C:Accession: JH0695; B45078  
R:Clark, J.A.; Deutch, A.Y.; Gallipoli, P.Z.; Amara, S.G.  
Neuron 9, 337-348, 1992  
A:Title: Functional expression and CNS distribution of a beta-alanine-sensitive neuronal  
A:Reference number: JH0695; MUID:92360310; PMID:1497897  
A:Accession: JH0695

A:Molecule type: mRNA  
A:Residues: 1-627 <CLN>  
A:Cross-references: GB:S42358; NID:g253466; PIDN:AA2850.1; PID:g253467  
R:Botden, L.A.; Smith, K.E.; Hartig, P.R.; Branchek, T.A.; Weinstank, R.L.  
J. Biol. Chem. 267, 21098-21104, 1992  
A:Title: Molecular heterogeneity of the gamma-aminobutyric acid (GABA) transport system.  
A:Reference number: A45078; MUID:93016029; PMID:1400419  
A:Accession: B45078

A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-627 <BOR>  
A:Cross-references: GB:M95763; NID:g202534; PIDN:AAA40607.1; PID:g202535  
A:Experimental source: brain  
A:Note: sequence extracted from NCBI backbone (NCBI:P116200)  
C:Superfamily: gamma-aminobutyric acid transporter  
C:Keywords: brain; glycoprotein; phosphoprotein; transmembrane protein

F:54-74/Domain: transmembrane #status predicted <TM2>  
F:81-101/Domain: transmembrane #status predicted <TM3>  
F:126-146/Domain: transmembrane #status predicted <TM3>  
F:223-239/Domain: transmembrane #status predicted <TM4>  
F:243-270/Domain: transmembrane #status predicted <TM5>  
F:297-317/Domain: transmembrane #status predicted <TM6>  
F:330-351/Domain: transmembrane #status predicted <TM7>  
F:383-403/Domain: transmembrane #status predicted <TM8>  
F:433-451/Domain: transmembrane #status predicted <TM9>  
F:468-488/Domain: transmembrane #status predicted <TM10>  
F:509-528/Domain: transmembrane #status predicted <TM11>  
F:548-566/Domain: transmembrane #status predicted <TM12>  
F:182,185,193/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:505,598/Binding site: phosphate (Thr) (covalent) #status predicted  
F:603/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 29.0%; Score 973; DB 2; Length 627;  
Best Local Similarity 35.8%; Pred. No. 7.2e-67;  
Matches 209; Conservative 109; Mismatches 236; Indels 30; Gaps 11;

Oy 22 SSQSLPPANKKALDNDITDLEAEP-----EAPPP-----ERWVNNNIEFLMSCIAT 67  
Db 2 TAEQALPLPNGKAAEARGSEALGGGGAAGTRARDKAVHERHNNKKEFLVSLVAGE 61  
Oy 68 SVGLGNVRRPFIATONGGAFVLYVIVLLVKGPPVYLLGECVLGQFSSRSNKKVW-SIS 126  
Db 62 IIGLGNVRRPFIATONGGAFVLYVIVLLVKGPPVYLLGECVLGQFSSRSNKKVW-SIS 121  
Oy 127 PAMKGTGYAQAAGCYILSYVYVIGLCLVYLLAMSFQATLPWALICQPEM--ENCVPSPDT 184  
Db 122 PLEEGIGATQVIEHNLNMYIILAMAFYLSNCFTELPWATGCHHEMTEKCYEFOKL 181  
Oy 185 LAASVNNIT--NGTSSAQLYFLRTVLOQSDGIEGGLAPITVLYLCLFIAMLVNGVVAR 242  
Db 182 NFNNSHVSLONNASPVMEFERVLAISDGE-HIGMLRELACILAAATTCYFCIKW 240  
Oy 243 GVKSGKAAYFLALFPYVMTITFTTIIILPGATGILFETVTPQAKLLLEGLVSAVTO 302  
Db 241 GTSKIGKVVYVYATPFIIMLILIRGVTLPQASGIRFYLYLPDLISRLSDQVWVDAQTO 300

Oy 303 VFEELVCTGPIIMFESSYNGRHRNIRYDAMIVTITDITFSLSCGTFIIGLNLAYELNS 362  
Db 301 IFESYATCLGCTLRALGSRNNYNNNCYRDCIMCLCNGSTSFVAGFAIFSLGPMAYDQV 360  
Oy 363 EVGDVAGAGTSLAFISYPPAIKATFQPOLESVLFPLMSVLAGISSVALLSTFNTLMD 422  
Db 361 PIAE-VAESGGLAFIYPAKAVTMTPLMTATLFFMLIFGLGDSQFCVCESLVAVVD 419  
Oy 423 AFPRV-----PPIYMAAMCSCGFLGLVYCTFGGQYILEVDHVGCT-FLVYFCAISE 475  
Db 420 MYPRFRRGRYRRELLIALSL-VSYFLGLVNLTEGGMVIFQLFDSYASGMCLEVAIFE 478  
Oy 476 LAGVFWIYGLNCLDIEFIMKRTGAYWRLCWGVIPATITVFFYFALLASNNLVFGDN 535  
Db 479 CVCIQWYYSNRFEDNIEDMIGYRPLSLIKMKVYVPGICAGIFIFIVKYRKLKNV 538  
Oy 536 VYPPAGYVSGYLMFLGMPFVPIGIFESLYKRYRFPSETIKK 579  
Db 539 YTPAMGVYIGLMLALMSMLCIPIMIFIKIKW-TEGTLPERLQK 581

## RESULT 11

choline transporter protein - rat  
S23431  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 20-Aug-1999  
A:Accession: S23431  
R:Mayser, W.; Schloss, P.; Betz, H.  
FEBS Lett. 305, 31-36, 1992  
A:Title: Primary structure and functional expression of a choline transporter express  
A:Reference number: S23431; MUID:92339519; PMID:1633856  
A:Accession: S23431

A:Molecule type: mRNA  
A:Residues: 1-635 <MAY>  
A:Cross-references: EMBL:X66494; NID:g55948; PIDN:CAA47119.1; PID:g55949  
C:Superfamily: gamma-aminobutyric acid transporter  
C:Keywords: glycoprotein; transmembrane protein  
F:61-83/Domain: transmembrane #status predicted <TM01>  
F:88-108/Domain: transmembrane #status predicted <TM02>  
F:133-160/Domain: transmembrane #status predicted <TM03>  
F:233-251/Domain: transmembrane #status predicted <TM04>  
F:260-277/Domain: transmembrane #status predicted <TM05>  
F:313-330/Domain: transmembrane #status predicted <TM06>  
F:342-366/Domain: transmembrane #status predicted <TM07>  
F:396-415/Domain: transmembrane #status predicted <TM08>  
F:445-462/Domain: transmembrane #status predicted <TM09>  
F:483-500/Domain: transmembrane #status predicted <TM10>  
F:521-542/Domain: transmembrane #status predicted <TM11>  
F:560-581/Domain: transmembrane #status predicted <TM12>  
F:192,197,548/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 29.0%; Score 972; DB 2; Length 635;  
Best Local Similarity 35.8%; Pred. No. 8.7e-67;  
Matches 209; Conservative 103; Mismatches 228; Indels 44; Gaps 11;

Oy 22 SSQSLPPANKKALDNDITDLEAEP-----EAPPP-----ERWVNNNIEFLMSCIAT 81  
Db 44 SSRALVPP-----RETWRQMDPIMSCGFAVGLGNVRRPFIAT 82  
Oy 82 YONGGAFVLYVIVLLVKGPPVYLLGECVLGQFSSRSNKKVW-SISPAKGTGYAQAAGC 141  
Db 83 YKNGGAFVLYVIVLLVKGPPVYLLGECVLGQFSSRSNKKVW-SISPAKGTGYAQAAGC 142  
Oy 142 YLSYVYVIGLCLVYLLAMSFQATLPWALICQPEM--NCVP---SDPLTAAVNNIT--- 193  
Db 143 YCNITVIVLWAMGFYIVLKSTTITLPAATCGHTNNTPPCVALFPHREDCAANSLANLTDQ 202  
Oy 194 --NGTSSAQLYFLRTVLOQSDGIEGGLAPITVLYLCLFIAMLVNGVVAR 251  
Db 203 LADRSPIVIERWENKVLRLSTGLEVP-CALNMEVATLCLACVAVLYFVWVGKSTKIV 261  
Oy 252 YFLALFPYVMTITFTTIIILPGATGILFETVTPQAKLLLEGLVSAVTO 311

Db 262 YFATFPVYVVLIVGVLLPGLDGIYYLKPDMSKLGSPQWIDAGTQIFESYAIGL 321  
Qy 312 GPIIMSSYNGERNIRYDAMIYTTLDFTSFGCTIFGLGNLAYLSEVGDVYAG 371  
Db 322 GALPALSSENNKCKDAIITLINSSTSEFAGFVFSILGEMAEQGVHISK-VAES 380  
Qy 372 GTSLAFISPDIAKTPQPOLFSVLFELMSVLGIGSSVALLSTFNLTAMDAFP-----R 426  
Db 381 GPGIAFIAPRAVYVLMVAPLMAALFFPMLLLGLDSGFVGEFTGLDLPLASYYFR 440  
Qy 427 VPVYMSAMTSCGFLGLGVYCTPGGQIILEVDHY--GGFVLVFCALISELAGVFWIYG 484  
Db 441 FQREISVALCCALCFVIDLMSVTDGMYFQLPFDYASASGFTL-LMQAFMECVVAVAYYG 499  
Qy 485 LENCLEIFEMLGKKTGAYWRLCWGVITPAIMTVEFYFALLASNNVYGVGDVYPTAGY 544  
Db 500 ADMRMDIACMTIGPCPCPMKMCWSEFTPLVCMGIFLPVNYVYKPLVYNNYVYPMGEA 559  
Qy 545 SGYLMLEFLGMTFVPIGIGFSLYKRYGTGFTSEITKKAHSKPSWG 588  
Db 560 MGMAFALSSMCLVPLHLGLGLLRK-GTMAERMQHL--TQPIWG 600

## RESULT 12

S19585  
serotonin transport protein - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 16-Feb-1997  
C:Accession: S19585  
R:Blakey, R.D.; Berson, H.E.; Fremeanu Jr., R.T.; Caron, M.G.; Peek, M.M.; Prince, H.K.;  
Nature 354, 66-70, 1991  
A:Title: Cloning and expression of a functional serotonin transporter from rat brain.  
A:Reference number: S19585; MUID:92049754; PMID:1944572  
A:Accession: S19585  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-607 <BLA>  
A:Cross-references: EMBL:X63253  
C:Superfamily: gamma-aminobutyric acid transporter  
C:Keywords: transmembrane protein

Query Match 29.0%; Score 971.5; DB 2; Length 607;  
Best Local Similarity 37.9%; Pred. No. 9e-67;  
Matches 222; Conservative 89; Mismatches 230; Indels 45; Gaps 14;  
Qy 26 SLPPANKKALNDIDDLAE--PPERWMSNNIEFLMSCIATSVGIGNWRPFPIAYQ 83  
Db 39 SIPAAT-----TTLVAEIRQGERETWGRKMDLFSVIGYAVDLGINMRPPICTYQ 88  
Qy 84 NGGGAFLVPYVILLVGRKPYVLECYLGOFSSRNSVYVW-SISPAKMGTYAAGCGY 142  
Db 89 NGGGAFLPPTIMAFPGIPLFYMELAGQYHRNGCISIMKICPIFGIGTALCIATFY 148  
Qy 143 ILSTYVVICGLCYLLAMSFOALPMAICQPEWE--NCVPSDPTLAASVNNIT--NCTS 197  
Db 149 IASYNNIIMALYLLISLTLRPLMTSCTNSMNTGNC-----TNYFQDNITWTLHSTS 203  
Qy 198 SAQILFRTVYQ--OSGIEGEGAPITMYLVLCFIFAMLMFVGVARVYKSSGRAAYLA 255  
Db 204 PAEEFYLRHVLIHQSGKLO-DIGTISWQTLCTVLTFTVYFISIMKVKISGRVWVTA 262  
Qy 256 LEPYVAVITLFTITLIPGATDGLIFEVTPQMAKILLEGVYSAVTOVFESLTCTGPRI 315  
Db 263 TFFPIVLSVILVRATILPGAMRGVYFLKPMWOKILLEGVAVDAAQIFESLGGFVLL 322  
Qy 316 MFSSYNGERNIRYDAMIYTTLDFTSFGCTIFGLGNLAYLSEVGDVYAGGTSL 375  
Db 323 AFASYNENNNCYODALVTSVYVNCMTSEFVSQVIFETVYGYAEKNEDVSEVADAGPSL 382  
Qy 376 AFISYPAIAKTPQPOLFSVLFELMSVLGIGSSVALLSTFNLTAMDAFPVPT----- 429  
Db 383 LFTIYAEAIAMNPASTFAIFELMLITLIGLSTFAGLEGVITAVLDEFPHIMAKRREWF 442

Qy 430 VYMSAMTSCGFLGLGVYCTPGGQIILEVDHYGFTFLVFCALISELAGVFWIYLEMJC 489  
Db 443 VLIYVITVLCSSLTL--TSGGAVVLLLEIATGPVALLVALLAANAASMTFTGIDQFC 499  
Qy 490 LDIEFMGKKTGAYWRLCWGVITPAIMTVEFYFALLASNNL-VFGDNVYPTAGVSGYL 548  
Db 500 SDVKEMLGFSQGWEMFRICQWVAISPLFLFICISFLMSPPQRLF--QYNPHMSVLCYC 557  
Qy 549 MLFLGMITVPIGIGFSLYKRYGTGFTSEITKKAHSKPSWGSRPSPE 594  
Db 558 IGMSSVICIPYIYIRLIS-TPGLKEKRIK-----SITPEPRE 596

## RESULT 13

G02095  
creatine transporter - human  
N:Alternate names: solute carrier family 6, member 8  
C:Species: Homo sapiens (man)  
C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 20-Aug-1999  
C:Accession: G02095  
R:Reichler, E.E.; Lu, F.; Muzny, D.; Gibbs, R.A.; Nelson, D.L.  
Submitted to the EMBL Data Library, September 1995  
A:Reference number: G09194  
A:Accession: G02095  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-635 <EUC>  
A:Cross-references: EMBL:U36341; NID:g1020318; PIDN:AAA79507.1; PID:g1020319  
C:Genetics:  
A:Gene: GDB:SLC6A8; CF-1  
A:Cross-references: GDB:434311; OMIM:600394  
A:Map position: Xq28-Xq28  
A:Intons: 88/1; 132/1; 215/2; 259/3; 304/3; 339/2; 381/1; 418/3; 464/3; 499/1; 532/3  
C:Superfamily: gamma-aminobutyric acid transporter

Query Match 29.0%; Score 971.5; DB 2; Length 635;  
Best Local Similarity 37.0%; Pred. No. 9.5e-67;  
Matches 206; Conservative 99; Mismatches 229; Indels 23; Gaps 10;  
Qy 49 PERWWSNNIEFLMSCIATSVGIGNWRPFPIAYONGGALVPYVILLVGRKPYVLE 108  
Db 50 PPRETWIRQMDIFMSCVGFAVGLGNWRPPIYLYKNGGVLPVLLVIALVAGIPIFLE 109  
Qy 109 CVLGOFSSRNSVYKWSISPAKMGTYAAGCGYILSYVVICGLCYLLAMSFOALPMP 168  
Db 110 ISLGOFMAGSINWNNICPLKRGISYASVIVFCNTIYIMLANGFYLVKSFYTLPLW 169  
Qy 169 AICQPEWE--NCVP---SDPTLAASVNNIT-----NGTSSAQLVFLFTVLOOSDIEGGL 218  
Db 170 ATCGTWTWTPCQVEIFRHEDCANSLAMLTCDQLADRSPIYEFENKVLRLSGLEVP- 228  
Qy 219 GAPIVYLVLCFIFAMLMFVGVARVYKSSGRAANFLAPYVVMITLFTITLIPGATDG 278  
Db 229 GALMNEVYLCLACVAVLFCVMKGVSKTGIVETAFEPVAVLVLLVRVLLPGALDG 288  
Qy 279 ILFETPQMAKILLEGVYSAVTOVFESLTCTGPRIIMFSSYNGERNIRYDAMIYTTLD 338  
Db 289 IITVTKPMSKLGSPQWIDAGTQIFESYAIGLGLTLAGSTNNRNNNCYDALLIALIN 348  
Qy 339 TETSLSGCTIFGLGNLAYLSEVGDVYAGGTSLAFISYPAIAKTPQPOLFSVLF 398  
Db 349 SGTSEFAFVYFESILGFMAADGVHISK-VAESGGGLAFIYAPRAVTLMPAPLMAALFF 407  
Qy 399 LMSVVLGIGSSVALLSTFNLTAMDAFP-----RPTVYMSAMTSCGFLGLGVYCTPGGQ 453  
Db 408 FMLLLIGLDSQFVGEFTGLDLPLASYYFRQREISVALCCALCFVIDLMSVTDGGM 467  
Qy 454 YILEVDHY--GGFVLVFCALISELAGVFWIYLEMJCLEFMGKKTGAYWRLCWGVI 511  
Db 468 YVQLEFDYTSASGTL-LMQAFMECVVAVAYYAGDRPDDIACMIGIYPCPMKMCWSEFF 526  
Qy 512 TPAIMTVEFYFALLASNNLVFGDNVYPTAGVSGYLMFLGMTFVPIGIGFSLYKRYGT 571

Db 527 TPLVCMGFIENNYVEEPLVNNNTYVPMWGMGAMFALSMCLVPLHLGLLRAK-G 585  
QY 572 TFESEIKKAFHSPKSWG 588  
Db 586 TMAERWQHL--TOPIMG 600

## RESULT 14

JC2386  
C:Species: Homo sapiens (man)  
C:Date: 28-Feb-1995 #sequence\_revision 05-Apr-1995 #text\_change 20-Aug-1999  
C:Accession: JC2386  
R:Scora, I.; Richman, J.; Santoro, G.; Wei, H.; Wang, Y.; Vanderah, T.; Horvath, R.; Nguy  
Blochem. Biophys. Res. Commun. 204, 419-427, 1994  
A:Title: The cloning and expression of a human creatine transporter.  
A:Reference number: JC2386; MUID:95032129; PMID:7945388  
A:Accession: JC2386  
A:Molecule type: mRNA  
A:Residues: 1-635 <SOR>  
A:Cross-references: GB:S74039; NID:9765233; PIDN:AAB32284.1; PID:9765234  
A:Experimental source: brain stem; spinal cord  
C:Superfamily: gamma-aminobutyric acid transporter  
C:Keywords: brain; glycoprotein; leucine zipper; phosphoprotein; spinal cord; transmembr  
F:62-82/Domain: transmembrane #status predicted <TM1>  
F:90-108/Domain: transmembrane #status predicted <TM2>  
F:134-160/Domain: transmembrane #status predicted <TM3>  
F:233-252/Domain: transmembrane #status predicted <TM4>  
F:263-277/Domain: transmembrane #status predicted <TM5>  
F:287-307/Region: leucine zipper motif  
F:314-329/Domain: transmembrane #status predicted <TM6>  
F:343-367/Domain: transmembrane #status predicted <TM7>  
F:400-418/Domain: transmembrane #status predicted <TM8>  
F:445-462/Domain: transmembrane #status predicted <TM9>  
F:482-499/Domain: transmembrane #status predicted <TM10>  
F:520-544/Domain: transmembrane #status predicted <TM11>  
F:563-583/Domain: transmembrane #status predicted <TM12>  
F:54/Binding site: phosphate (Thr) (covalent) (by CAMP-dependent kinase) #status predict  
F:192,197,548/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:256/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status predic

Query Match 29.0%; Score 971; DB 2; Length 635;  
Best Local Similarity 35.2%; Pred. No. 1e-66;  
Matches 206; Conservative 104; Mismatches 229; Indels 46; Gaps 10;  
QY 22 SSOISLPANNAKALNDITDLEAPPERMWNNEEFMLSCATSVGLGNVRPPIA 81  
Db 44 SSRALVPP-----RETWTROMDFIMSCVGFAGLGNVRPFLC 82  
QY 82 YONGGAFLLVYVIVLLVGRKPYVLECVLQFSSRSVSKVMSISPAKGTGYAQAACG 141  
Db 83 YKNGGVFLIPYVIALVGVGIPFLFELISLGQFMKAGSINVMNCPLEFGKGYASMTVF 142  
QY 142 YIISGVVIGLCICLYLAMSQATLPAICQPEW-----ENCVSDDPLASVN 190  
Db 143 YCNYYIVLWAMGFYVVKSFETTLPMATGHTWNTDCEVEIFRHECCA-NDSLANTCD 201  
QY 191 NITNGTSSADLYELFTVLOOSDIEGIGLAPVIVLCLFTAMLVGVAVARGKSSGA 250  
Db 202 QLANDRSPVIEFWENKULRLSTGLEVP-GALNMEVITLCLACWLVVFCWKYKSGIKI 260  
QY 251 AYFLALFPYVMTLFTTIIILPGANDGILFVTPQAKLLELGVMSAVTQVFSLTVC 310  
Db 261 VYFATPFPYVVLVLRVGLVLPALDGIITLKPMSKLGSPQVMDAGTQIFFSYAI 320  
QY 311 TGPIIMSSYNGFHNITRDAMVITTDFTSFLSGTIFGILGNLALYELNSEKDVYGA 370  
Db 321 LGALTALGSYNRENNQYKAIILALINSSTSEFAGVVSIIIGFMAETEGVHLSK-VAE 379  
QY 371 GGTSLAFISYPDAIAKTFQPOLFSVLEFFLMMSVGLSSVALLSTFTMLMDAPF----- 425  
Db 380 SGGLATIAIAPRAVTLMPVAPLWALFFPMLLLGLDSDQVGVGSGFITGLDILLPASYIF 439

QY 426 RPTVYMSAMTCSCGELLGLVYCTPGGOYIILEVDHY--GGTFLVLCALSELAGVFYFI 483  
Db 440 RFOREISVALCALCEFTLDLSMTDGGKTYVQLDYIASAGTTL-LWQAEVCYVAANY 498  
QY 484 GLENCLDIEPMLGKKTGAYWRLCGVITPAIMTTFVFPALLASNNLFGDNNYPTAGY 543  
Db 499 GADRFMDIACMIGRCPMWMKWCMSFTPLVCGIFLNFVNYEPLVNNNTYYPMMGE 558  
QY 544 VSGTLMFLGTFPPIGIFESLYKRYGTSEETTKKAFHSPKSWG 588  
Db 559 AMGNAFALSMCLVPLHLGLLRAK-GTMAERWQHL--TOPIMG 600

## RESULT 15

A46061  
Na(+)-dependent creatine transporter - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 12-Apr-1995  
C:Accession: A46061  
R:Guibal, C.; Kilmann, M.W.  
J. Biol. Chem. 268, 8418-8421, 1993  
A:Title: A Na(+)-dependent creatine transporter in rabbit brain, muscle, heart, and k  
A:Reference number: A46061; MUID:93231984; PMID:8473283  
A:Accession: A46061  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-635 <GUI>  
A:Experimental source: brain and muscle  
A:Note: sequence inconsistent with the nucleotide translation  
A:Note: sequence extracted from NCBI backbone (NCBI:129571, NCBI:129580)  
C:Superfamily: gamma-aminobutyric acid transporter

Query Match 28.9%; Score 970.5; DB 2; Length 635;  
Best Local Similarity 36.8%; Pred. No. 1.1e-66;  
Matches 205; Conservative 97; Mismatches 232; Indels 23; Gaps 10;  
QY 49 PERMWSNIEFLMSCATSVGLGNVRPPIAYONGGAFLLVYVIVLLVGRKPYVLE 108  
Db 50 PRRTWTRMDFIMSCVGFAGLGNVRPFLCCKNGGVFLIYVIALVGVGIPFLF 109  
QY 109 CVLQFSSRSVSKVMSISPAKGTGYAQAACGIIISYVIGLCILYLAASFQATLPW 168  
Db 110 ISLQGFMAKAGSINVMNCPLEFGKGYASMTVFYCNYYIVLWAMGFYVVKSFETTL 169  
QY 169 AICQPEW--NCVP---SDPTLASVNNIT-----NGTSSQLLEFTVLOOSGIGGL 218  
Db 170 ATCGHTWNTDCEVEIFRHEDCANGSLANTCDQLAERRSPVIEFWENKVLNLSGLEVP- 228  
QY 219 GAPITVYVLCFLIMLWVGVAVARGVSKRAAYFLALFPYVMTLFTTIIILPGATDG 278  
Db 229 GALNMEVITLCLACWLVVFCWKYKSGIKIYFTAPFVVLVLRVGLLPGALDG 288  
QY 279 IIFVTPQAKLLELGVMSAVTQVFSLYVCTPIIMFSSYNGFRHNITRDAMIVTTLD 338  
Db 289 IITYLKPDWSKLSRPQVMDAGTQIFFSYALGIALALAGSYNPNKCYKDAIILAN 348  
QY 339 TFTSFLSGCTIFGILGNLALYELNSEVDVAGGTSIAFISYPAIKTROPOLFSVLEF 398  
Db 349 SGTFEFAGFVFSITLGMATBOGVHISK-VAESRPGALFIAPRAVTLMPAPLMAALFF 407  
QY 399 LMSVGLIGSSVALLSTFTMLMDAPF----RPTVYMSAMTCSCGELLGLVYCTPGGO 453  
Db 408 FMLLLGLDSDQFVGEVETIGLIDLLPASYIFRQREISVALCALCEFTLDLSMTDGGM 467  
QY 454 YIILEVDHY--GGTFLVLCALSELAGVFYFIILEVDHY--GGTFLVLCALSELAGVFYFI 511  
Db 468 YVQLFDYYSASGTTL-LWQAEVCYVAANYGADRFMDIACMIGRCPMWMKWCMSF 526  
QY 512 TPAIMTTFVFPALLASNNLFGDNNYPTAGVYGYLMLFGMTFVPIGIGFSLYKYRTG 571  
Db 527 TPLVCMGFIENNYVEEPLVNNNTYVPMWGMGAMFALSMCLVPLHLGLLRAK-G 585

Mon Jul 28 08:01:35 2003

us-09-991-458-2.rpr

Page 9

OY 572 TFSSTIKKAFHSKPSWG 588  
| : | : : | |  
Db 586 TMAERWOHL-TOPTWG 600

Search completed: July 21, 2003, 09:26:11  
Job time : 45 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2003, 09:24:45 ; Search time 24 Seconds

(without alignments)  
1093.938 Million cell updates/sec

Title: US-09-991-458-2

Perfect score: 3354

Sequence: 1 MNDQVNGGFESESEPKMEPK.....TSRVKHLWYSITGAYRRNIN 633

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1080.5	32.2	797	1	SGA5_HUMAN
2	1076.5	32.1	797	1	SGA5_RAT
3	1027	30.6	637	1	SGA7_RAT
4	1020	30.4	636	1	SGA7_HUMAN
5	1011	30.1	599	1	SGA1_HUMAN
6	1008	30.1	599	1	SGA1_RAT
7	998	29.8	630	1	SGA4_MACMU
8	992.5	29.6	598	1	SGA1_MUSCO
9	991	29.5	630	1	SGA4_HUMAN
10	991	29.5	630	1	SGA4_MOUSE
11	987	29.4	630	1	SGA4_BOVIN
12	981	29.2	630	1	SGA4_RAT
13	979	29.2	614	1	SGA6_RAT
14	978.5	29.2	598	1	SGA1_MOUSE
15	978.5	29.2	632	1	SGA1_HUMAN
16	975	29.1	635	1	SGA8_BOVIN
17	974.5	29.1	620	1	SGA6_BOVIN
18	973	29.0	627	1	SGA6_RAT
19	972	29.0	635	1	SGA8_RAT
20	971.5	29.0	635	1	SGA8_HUMAN
21	971	29.0	622	1	SGA4_DROME
22	970	28.9	630	1	SGA4_CAVPO
23	968.5	28.9	620	1	SGA6_HUMAN
24	967	28.8	620	1	SGA9_HUMAN
25	965.5	28.8	620	1	SGA6_CANFA
26	965	28.8	602	1	SGA6_RAT
27	964	28.7	614	1	SGA6_MOUSE
28	963	28.7	627	1	SGA6_MOUSE
29	959	28.6	602	1	SGA6_MOUSE
30	953.5	28.4	621	1	SGA6_RAT
31	949.5	28.3	620	1	SGA3_HUMAN
32	949.5	28.3	635	1	SGA8_RABIT
33	943.5	28.1	621	1	SGA6_MUSCO

34	940.5	28.0	614	1	SGA6_CANFA	P27799	canis fam1
35	939	28.0	619	1	SGA3_MOUSE	O61327	mus musculu
36	936	27.9	615	1	SGA2_BOVIN	P51143	bos taurus
37	936	27.9	638	1	SGA9_BOVIN	O28039	bos taurus
38	935	27.9	633	1	SGA9_MOUSE	P28571	mus musculu
39	935	27.9	633	1	SGA9_RAT	P28572	rattus norv
40	934.5	27.9	621	1	SGA6_MOUSE	O35316	mus musculu
41	933	27.8	619	1	SGA3_RAT	P23977	rattus norv
42	932.5	27.8	614	1	SGA6_HUMAN	P48065	homo sapien
43	931.5	27.8	614	1	SGA6_RABIT	P48055	oryctolagus
44	930	27.7	617	1	SGA2_MOUSE	O55192	mus musculu
45	920.5	27.4	617	1	SGA2_HUMAN	P23975	homo sapien

## ALIGNMENTS

RESULT 1	ID	SEQUENCE	STANDARD	PRT	797 AA
AC	SGA5_HUMAN	O9Y345	O9Y288; O9BX77;		
DT	16-OCT-2001	(Rel. 40, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Sodium- and chloride-dependent glycine transporter 2 (GLYT2) (GLYT-2).				
GN	SICGA5 OR GLYT2.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OX	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
NCBI_TaxID	9606;				
[1]	SEQUENCE FROM N.A.				
RX	MEDLINE-99059482; PubMed-9845349;				
RA	Morrow J.A., Collie I.T., Dunbar D.R., Walker G.B., Shahid M.,				
RA	Hill D.R.;				
RT	"Molecular cloning and functional expression of the human glycine				
RT	transporter GLYT2 and chromosomal localisation of the gene in the				
RT	human genome.";				
FEBS Lett.	439:334-340(1998).				
[2]	SEQUENCE FROM N.A.				
RP	TISSUE-Spinal cord;				
RX	MEDLINE-99310916; PubMed-10381548;				
RA	Gallagher M.J., Burgess L.H., Brunden K.R.;				
RT	"Characterization of multiple forms of the human glycine transporter				
RT	type-2.";				
RL	Brain Res. Mol. Brain Res. 70:101-115(1999).				
[3]	SEQUENCE FROM N.A.				
RP	TISSUE-Spinal cord;				
RT	Luyten W.;				
RT	"Cloning and expression of a human glycine transporter type II.";				
RL	Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.				
[4]	SEQUENCE FROM N.A., AND VARIANTS ARG-184; ASN-463 AND ALA-751.				
RP	TISSUE-Spinal cord;				
RX	MEDLINE-20074920; PubMed-10606742;				
RA	Evans J., Herdon H., Cairns W., O'Brien E., Chapman C., Terrett J.,				
RA	Gloger I.;				
RT	"Cloning, functional characterisation and population analysis of a				
RT	variant form of the human glycine type 2 transporter.";				
FEBS Lett.	463:301-306(1999).				
CC	- FUNCTION: Terminates the action of glycine by its high affinity				
CC	sodium-dependent reuptake into presynaptic terminals. May be				
CC	responsible for the termination of neurotransmission at				
CC	syncholine-sensitive glycinergic synapses.				
CC	- SUBCELLULAR LOCATION: Integral membrane protein.				
CC	- TISSUE SPECIFICITY: Expressed in medulla, and to a lesser extent				
CC	in spinal cord and cerebellum.				
CC	- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER				
CC	FAMILY (SNF).				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ksb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: AF085412; AAC95145.1; -  
 DR EMBL: AF142501; AAD27892.1; -  
 DR EMBL: AF117999; AAK12641.1; -  
 DR EMBL: AF352733; AAK29670.1; -  
 DR Genew: HGNC:11051; SLC6A5.  
 DR MIM: 604159; -  
 DR InterPro: IPR000175; Na/ntran\_symport.  
 DR Pfam: PF00209; SNF; 1.  
 DR ProDom: PD000448; Na/ntran\_symport; 2.  
 DR PROSITE: PS00610; NA\_NEUROTAN\_SYM\_1; 1.  
 DR PROSITE: PS00734; NA\_NEUROTAN\_SYM\_2; 1.  
 DR PROSITE: PS50267; NA\_NEUROTAN\_SYM\_3; 1.  
 DR Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;  
 KW Symport; Polymorphism.

FT DOMAIN 1 199  
 FT TRANSMEM 200 220  
 FT TRANSMEM 228 247  
 FT TRANSMEM 271 291  
 FT DOMAIN 292 393  
 FT TRANSMEM 394 412  
 FT TRANSMEM 421 438  
 FT TRANSMEM 474 491  
 FT TRANSMEM 503 524  
 FT TRANSMEM 557 576  
 FT TRANSMEM 604 622  
 FT TRANSMEM 638 658  
 FT TRANSMEM 679 698  
 FT TRANSMEM 717 735  
 FT DOMAIN 737 797  
 FT CARBOHYD 343 343  
 FT CARBOHYD 353 353  
 FT CARBOHYD 358 358  
 FT CARBOHYD 364 364  
 FT VARIANT 184 184  
 FT VARIANT 463 463  
 FT VARIANT 751 751  
 FT VARIANT 751 751  
 FT CONFLICT 24 24  
 FT CONFLICT 102 102  
 FT CONFLICT 155 155  
 FT CONFLICT 188 188  
 FT CONFLICT 362 362  
 FT CONFLICT 582 582  
 SQ SEQUENCE 797 AA; 87359 MW; BBA089AB7EBBDA CRC64;

Query Match 32.28; Score 1080.5; DB 1; Length 797;  
 Best Local Similarity 36.28; Pred. No. 7.2e-64;  
 Matches 221; Conservative 106; Mismatches 213; Indels 39; Gaps 10;

QY 54 MSNIEFLMCIATSVGIGNWRPPTAYONGGAGFAPYIVLLVGRPYVLECYGQ 113  
 DB 194 WSKSLDFILSVNGAVGLGNWRPPTAYONGGAGFAPYIVLLVGRPYVLECYGQ 253  
 QY 114 FSSRSNVKWSISIPAMKGTGYAAGCGYILSYVYVIGLCLLYLANSFOATLPAICOP 173  
 DB 254 FASGQPSVWKAIPALOGCGIAMLIIIVLAIYVNVIICTLFYLFASVFLWPGSCNN 313  
 QY 174 EW-----ENCVPSD-PTL-----AASVNNIT--NGISSA-----QL 201  
 DB 314 PMNTECKDKTKLLLDSCVSDHDKIKNSTFCMTAYPNVTWNFTSQANKTEVSGSEE 373  
 QY 202 YFLRTVLOQSDIGEGGAPITWYLVLCIFETAMLWFGVANGVSSGKAAYFLALFPYV 261  
 DB 374 YFKFVLIKISAGIEYR-GEIRWPLALCLFLAWLVIVASLAKGITSKVYVITFTFFYV 432

QY 262 MIFLFTITLLPGATDGIIFPVTPONAKLLELVWYSAVOVEFSITVCTPIIMFSYN 321  
 DB 433 LVILLRGVLLPAGAGACITWFTPKMKTLDATVWMDAQIFSSAMGGLITLSSYN 492  
 QY 322 GFRNNIYRDAMIYITLDTFTSLSGCTIFGILGNLYELNSVGDVAGAGTSLATISP 361  
 DB 493 KFRNNCYRDLIYCTNSATSIKFAFVIFSVIGFMANERKVIENVADQ-GPGIARFVYP 551  
 QY 382 DATAKTPQDPSVLPFLMSVIGISGVALISTFNPLANDAPRPPTYMSMT---CS 438  
 DB 552 EALTPLSPFWMIIFFLMLTLTGDTMEFTIETIVTISIDEPKYLRTKHPVFTLGCCT 611  
 QY 439 CGFLGALVYCTPGGOVYLLVDHYGTFVLFCALISLACVFNWYIGENLCLDIEMPLGK 498  
 DB 612 CFTIMGPMITOGGIVFQVLDVRYAASVALYIAIFELVIGISYVGLQRCEDEIMWIGF 671  
 QY 499 KTGAYRFLCNGVITPAIMTTFVFFALLASNNLVFGDNYVPTGYVSGYLMELGKTFP 558  
 DB 672 QPNIFWCVKAWFVPTLFTFLFCFSFYQWPMYTG-SYRYPNMSMYLWMLACSVIWP 730  
 QY 559 IGIGFSLKYKRTGFSEETIKAFHSKPSMGPRSPRRRE 597  
 DB 731 IMFYIKMH-LAPGRFLERLKVCSPPDWPGLAQRGE 768

RESULT 2  
 ID 56A5\_RAT STANDARD; PRT; 799 AA.  
 AC P88295;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Sodium- and chloride-dependent glycine transporter 2 (GlyT2) (GlyT-2).  
 GN SLC6A5 OR GLYT2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=94043045; PubMed=8226790.  
 RA Liu Q.-R., Lopez-Corcuera B., Mandlyan S., Nelson H., Nelson N.;  
 RT "Cloning and expression of a spinal cord- and brain-specific glycine  
 RT transporter with novel structural features.";  
 RL J. Biol. Chem. 268:22802-22808(1993).  
 CC -!- FUNCTION: Terminates the action of glycine by its high affinity  
 CC sodium-dependent reuptake into presynaptic terminals. May be  
 CC responsible for the termination of a spinal cord- and brain-specific glycine  
 CC strychnine-sensitive glycinergic synapses.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: Specifically expressed in spinal cord, brain  
 CC stem, and to a lesser extent in the cerebellum.  
 CC -!- SIMILARITY: BELONGS TO THE SODIUM-NEUROTANSMITTER SYMPORTER  
 CC FAMILY (SNF).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ksb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: L21672; -; NOT\_ANNOTATED\_CDS.  
 DR InterPro: IPR000175; Na/ntran\_symport.  
 DR Pfam: PF00209; SNF; 1.  
 DR PRINTS: PR00176; NANEUSPORT.  
 DR PROSITE: PS00610; NA\_NEUROTAN\_SYM\_1; 1.  
 DR PROSITE: PS00734; NA\_NEUROTAN\_SYM\_2; 1.  
 DR PROSITE: PS50267; NA\_NEUROTAN\_SYM\_3; 1.  
 KW Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;  
 KW Symport.



```

FT DOMAIN 1 201 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 202 222 1 (POTENTIAL).
FT TRANSMEM 230 249 2 (POTENTIAL).
FT TRANSMEM 273 293 3 (POTENTIAL).
FT DOMAIN 294 395 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 396 414 4 (POTENTIAL).
FT TRANSMEM 423 440 5 (POTENTIAL).
FT TRANSMEM 476 493 6 (POTENTIAL).
FT TRANSMEM 505 526 7 (POTENTIAL).
FT TRANSMEM 559 578 8 (POTENTIAL).
FT TRANSMEM 606 624 9 (POTENTIAL).
FT TRANSMEM 640 660 10 (POTENTIAL).
FT TRANSMEM 681 700 11 (POTENTIAL).
FT TRANSMEM 719 737 12 (POTENTIAL).
FT DOMAIN 739 799 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 366 366 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 799 AA: 87908 MW: 9DBE239A7B071FCD CRC64;

```

## Query Match

Best Local Similarity 32.1%; Score 1076.5; DB 1; Length 799;

Matches 224; Conservative 107; Mismatches 228; Indels 39; Gaps 10;

```

QY 35 ALNDIDTDLAEPPERWWSNNIEFLMCIATSVGLGNWFFPIAYONGGAFVLPYV 94
DB 177 ATTITPEDEGDEGNARGNMSSKDFIISMVGYANGLVNWFPIAYONGGAFVLPYV 236
QY 95 IVLLVGRPVYILECVLQFSSRSNYSWISPAKMGTYQAAGCYLLSYVYVIGCIG 154
DB 237 MMLAAGLPPIEFLESLGQFASQGVSWKAIPALQCGIAMIISVLAIVYNYICYT 296
QY 155 LYVLAISPOATLPMALICQPEW-----ENCVPSD-PTL-----ASVNNI 192
DB 297 LEYLFASFVSVLPMGSCNNPMPTPECKDKTKLLDSVYGDHPKIQIKNSTFCMTAYPM 356
QY 193 T--NCTSSA-----OLYFLRTVLOOSDGIIEGLGAPVWLVLCLFAMLVFQVVAR 242
DB 357 TMVNTSQANKTFVSGSEYEFKYLKISAGIEYF-GELRMWLPICLFIAMVIVYASLAK 415
QY 243 GYKSSGKAYFLALPPYVMTLFTITLLPGATDGLIEFVTPQNAKLELCVWSAVTQ 302
DB 416 GIKTSGKVVYFATPPYVVLVILIRGYVLPAGAGIMWFITPKMEKLDATVMDAATQ 475
QY 303 VFEFLVTCGPIIMSSYNGFRHNYRDAMVITLDTFTLSGCTIRIILGNAYELS 362
DB 476 IFFSLAAMGGLITLSTINKFHNCCYRDLVITCTNSATSIAGFVIESVIGFMANERKV 535
QY 363 EYGDVYVAGGTSIAFISYPAIAKTFQDQLFSLVFLMMSVGLIGSSVALLSTFNTIAM 422
DB 536 NIENNVADQ-GPGIAVYVPEALTRPLSPFMAIFFMALLTGLDMETIETIYISID 594
QY 423 AFRPVYVYMSAMT---CSCGFLGLVYCTPGQYILLVHVGCTFVLFCALISLAV 479
DB 595 EEPKYLIRHKPVFTLGCICCFEIMGFPMTOGGIMFOLVDVYAAISYALIAIFELVGI 654
QY 480 FVIYGLNLCIDIEEMLKKTGAYWRICGVITPAMTVFFYVALLASNNLVFGDNYVP 539
DB 655 STVYGLQRPCEDEIEMNIGFOPNIEFWKVCVAFVPTLIFLILCFESQWMEPMYTG-SYRP 713
QY 540 TAGVYSGYLMFLGTFFVIGIGISLYKRTGTSETIKAKHSRPMGPRPRERE 597
DB 714 NMSVNLGMLMACSYIWMIPMVIKMY-LAPGRFIERLKLVCSPQDMGPFLLAQHGE 770

```

## RESULT 3

```

66A7_RAT STANDARD: PRT: 637 AA.
AC P28573:
DT 01-DEC-1992 (rel. 24, Created)
DT 16-OCT-2001 (rel. 40, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)

```

```

DE Sodium-dependent proline transporter.
GN SLC6A7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RC MEDLINE=92265305; PubMed=1350201;
RA Fremieu R.T., Jr., Caron M.G., Blakey R.D.;
RT "Molecular cloning and expression of a high affinity L-proline
RT transporter expressed in putative glutamatergic pathways of rat
RT brain."
RL Neuron 8:915-926(1992).
CC -1 FUNCTION: TERMINATES THE ACTION OF PROLINE BY ITS HIGH AFFINITY
CC SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
CC -1 SUBCELLULAR LOCATION: Integral membrane protein.
CC -1 TISSUE SPECIFICITY: EXPRESSED IN SUBPOPULATIONS OF PUTATIVE
CC GLUTAMATERGIC PATHWAYS OF RAT BRAIN.
CC -1 SIMILARITY: BELONGS TO THE SODIUM-NEUROTRANSMITTER SYMPORTER
CC FAMILY (SNP).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M8811; AAA41541.1; ALT TERM.
CC Interpro: IPR000175; Na/nttran_symport.
CC Pfam: PF00209; SNF. 1.
CC PRINTS: PR00176; NANEUSMPORT.
CC PRODOM: PD000448; Na/nttran_symport; 1.
CC PROSITE: PS00610; NA_NEUOTRAN_SYM_1; 1.
CC PROSITE: PS00754; NA_NEUOTRAN_SYM_2; 1.
CC PROSITE: PS50267; NA_NEUOTRAN_SYM_3; 1.
CC KW Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
CC Symport; Amino-acid transport.
CC -----
CC DOMAIN 1 45 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 46 66 1 (POTENTIAL).
CC TRANSMEM 74 93 2 (POTENTIAL).
CC TRANSMEM 117 137 3 (POTENTIAL).
CC DOMAIN 138 214 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 215 223 4 (POTENTIAL).
CC TRANSMEM 242 259 5 (POTENTIAL).
CC TRANSMEM 295 312 6 (POTENTIAL).
CC TRANSMEM 324 345 7 (POTENTIAL).
CC TRANSMEM 378 397 8 (POTENTIAL).
CC TRANSMEM 425 443 9 (POTENTIAL).
CC TRANSMEM 459 479 10 (POTENTIAL).
CC TRANSMEM 500 519 11 (POTENTIAL).
CC TRANSMEM 538 556 12 (POTENTIAL).
CC DOMAIN 557 637 CYTOPLASMIC (POTENTIAL).
CC CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 637 AA: 71090 MW: 9627E6D5B8C9408 CRC64;

```

## Query Match

Best Local Similarity 30.6%; Score 1027; DB 1; Length 637;

Matches 227; Conservative 90; Mismatches 236; Indels 18; Gaps 11;

```

QY 41 DDDLEAF-PPERWWSNNIEFLMCIATSVGLGNWFFPIAYONGGAFVLPYVILL 99
DB 26 DVDLDVDFADAGNMTGKLDPLSLCIGYVGLGNWVFPYRATNNGGAFVLPYVLMAL 85
QY 100 VGRPVYILECVLQFSSRSNYSWISPAKMGTYQAAGCGYLLSYVYVIGLCIAYLIA 159
DB 86 CGPLPFLFELSGLQFSSLSGLPPLAVWKISPLFKGAGAMLLIVGVALIYVMIILAVLYLF 145
QY 160 MSFQATLPMALICQPEW-ENCY---PSDPTLAASVNNITNGTSSAQLYFLRTVL--QQS 211

```

```
Dh 146 ASLTSNLPWEHCMMWNTERCLERHPKDGNGALPL-NLSTVSPSEBYWSRYVLIHIGS 204
Qy 212 DGIIGGLGAPIMYLVLCFLIAMLMVFGVARGVSSGKAAYELAPFYVWITLFTITII 271
Dh 205 QGI-GSGEIRWNLCCLLAMVIFELCIKGVSSGVVYFTATFPYLILMLLVGVT 263
Qy 272 LPGAIDGLFPVTPQMAKLELGVYSAVQVFFSLVYCTGPITMSSSYNGFRHNYRDA 331
Dh 264 LPGAAMKGIQFYLTPQFHLHLLSSKWIEMALQIFKSLGVFGGLLTFASYNFHONIVRDT 323
Qy 332 WIYVTLDTFSLGCTIFGLGNLAAYLSEVDVAGAGTSLAFISYPAIAKTRPOQ 391
Dh 324 FIVLGNATISILGFAIFSVLGYMSQELGVPV-DQAKAPGGLAFVYTPQAMMLPLSP 382
Qy 392 LFSVLEFLMSVILGSSVALLSTFNTLAMDAPF--RVPTVYSAMTSCGFLGLGYVC 448
Dh 383 FMSLFFFMILLTLGDSQFALETIVAVDEFPYLRPKRAVSGSLICVAMYLMGLILT 442
Qy 449 TPGQGYLLEVDHNGGFVLFCALISELAGVFWIYGLLENCLDIEFMLGKKTGAYWRLCW 508
Dh 443 TDGGMVYLVLLDDYSASFGMLVAVVITCLAVTRYGIQRCDIHMLGFKPGLYFRACW 502
Qy 509 GVTPAINTVYEFALLASNNLVFGDNVYPTAGVSGYMLFLGMTFVPYIGIGFSLYKY 568
Dh 503 LFSPLALLLVISYKYPSEYG-SYRPPAMAEILGILMGLSLCMTIFAGMLVAVLR- 560
Qy 569 RTGTFSETIKKAFHKSPPGPRSPREREM 599
Dh 561 EEGSLMERLQASRPAMDGP-SLEENRTGM 590

RESULT 4
56A7_HUMAN STANDARD; PRT; 636 AA.
AC 099884;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sodium-dependent proline transporter.
GN SLC6A7 OR PROT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hippocampus;
RX MEDLINE=95379780; PubMed=7651355;
RA Shafqat S., Velaz-Palacio M., Henzl V.A., Whitney K.D.,
RA Yang-Feng T.L., Seidlin M.F., Fremean R.T. Jr.
RT "Human brain-specific L-proline transporter: molecular cloning,
RT functional expression, and chromosomal localization of the gene in
RT human and mouse genomes."
RL Mol. Pharmacol. 48:219-229(1995).
CC -1- FUNCTION: TERMINATES THE ACTION OF PROLINE BY ITS HIGH AFFINITY
CC SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: BRAIN.
CC -1- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC FAMILY (SNF).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S80071; ABA47007.2; -
DR Genew; HGNC:11054; SLC6A7.
DR MIM; 606205; -
DR Interpro; IPR000175; Na/ntran_symport.
```

```
DR Pfam; PF00209; SNF; 1.
DR ProDom; PD000448; Na/ntran_symport; 1.
DR PROSITE; PS00610; NA_NEUTROTAN_SYM_1; 1.
DR PROSITE; PS00754; NA_NEUTROTAN_SYM_2; 1.
DR PROSITE; PS0267; NA_NEUTROTAN_SYM_3; 1.
DR Neurotransmitter transporter; Transport; Transmembrane; Glycoprotein;
KW Symport; Amino-acid transport; Polymorphism.
DR DOMAIN 1 45
FT TRANSMEM 46 66 1 (POTENTIAL).
FT TRANSMEM 74 93 2 (POTENTIAL).
FT TRANSMEM 117 137 3 (POTENTIAL).
FT DOMAIN 138 214 4 (POTENTIAL).
FT TRANSMEM 215 233 5 (POTENTIAL).
FT TRANSMEM 242 259 6 (POTENTIAL).
FT TRANSMEM 295 312 7 (POTENTIAL).
FT TRANSMEM 324 345 8 (POTENTIAL).
FT TRANSMEM 378 397 9 (POTENTIAL).
FT TRANSMEM 425 443 10 (POTENTIAL).
FT TRANSMEM 459 479 11 (POTENTIAL).
FT TRANSMEM 500 519 12 (POTENTIAL).
FT TRANSMEM 538 556 12 (POTENTIAL).
FT DOMAIN 557 636 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARIANT 345 345 L->V (IN DBSNP:1468564).
FT /FTID=VAR_011390.
SQ SEQUENCE 636 AA; 70892 MM; 9A2D1071P59E409B CRC64;
Query Match 30.4%; Score 1020; DB 1; Length 636;
Best Local Similarity 39.8%; Pred. No. 5.4e-60;
Matches 227; Conservative 87; Mismatches 239; Indels 18; Gaps 10;
Qy 41 DTDLEAE-PPRWVWNNIEFLMSCIATSVGLGVNWRFPPTAYONGGAFVLPYVILL 99
Dh 26 DVDLDVDFPAHRGMWTKIDFLSLCIGCVGLGVNWRFPYRPAVINGGAFVLPYVILLAI 85
Qy 100 VGRVYVLECYLGPFSSRSNYSKWSISPAKGTGTAQAGCGYLLSYVIGCLYLA 159
Dh 86 CGIPFLFLESLGPFSSIGPLAVKISPLFGAAMLLVGLVAIYNNMIAVLYELF 145
Qy 160 MSFOATLPMAICQPEM-----ENCYPSDPTLASVNNITGSTSAOLYFRVL--QOS 211
Dh 146 ASLTSNLPWEHCMMWNTERCLERHRSKDGNGALPL-NLCTVSPSEBYWSRYVLIHIGS 204
Qy 212 DGIIGGLGAPIMYLVLCFLIAMLMVFGVARGVSSGKAAYELAPFYVWITLFTITII 271
Dh 205 QGI-GSGEIRWNLCCLLAMVIFELCIKGVSSGVVYFTATFPYLILMLLVGVT 263
Qy 272 LPGAIDGLFPVTPQMAKLELGVYSAVQVFFSLVYCTGPITMSSSYNGFRHNYRDA 331
Dh 264 LPGAAMKGIQFYLTPQFHLHLLSSKWIEMALQIFKSLGVFGGLLTFASYNFHONIVRDT 323
Qy 332 WIYVTLDTFSLGCTIFGLGNLAAYLSEVDVAGAGTSLAFISYPAIAKTRPOQ 391
Dh 324 FIVLGNATISILGFAIFSVLGYMSQELGVPV-DQAKAPGGLAFVYTPQAMMLPLSP 382
Qy 392 LFSVLEFLMSVILGSSVALLSTFNTLAMDAPF--RVPTVYSAMTSCGFLGLGYVC 448
Dh 383 FMSLFFFMILLTLGDSQFALETIVAVDEFPYLRPKRAVSGSLICVAMYLMGLILT 442
Qy 449 TPGQGYLLEVDHNGGFVLFCALISELAGVFWIYGLLENCLDIEFMLGKKTGAYWRLCW 508
Dh 443 TDGGMVYLVLLDDYSASFGMLVAVVITCLAVTRYGIQRCDIHMLGFKPGLYFRACW 502
Qy 509 GVTPAINTVYEFALLASNNLVFGDNVYPTAGVSGYMLFLGMTFVPYIGIGFSLYKY 568
Dh 503 LFSPLALLLVISYKYPSEYG-SYRPPAMAEILGILMGLSLCMTIFAGMLVAVLR- 560
Qy 569 RTGTFSETIKKAFHKSPPGPRSPREREM 599
Dh 561 EEGSLMERLQASRPAMDGP-SLEENRTGM 590

RESULT 5
```

Db	10	DGQI-----SPEVSEAPVANDKPKTLVAVKQKKAADLPDRDTWGKREDFL	54
Qy	62	MSCIATSVGIGVNRFPPIAIYQNGCGAFVPPYVVLVLKRPVYIECVLGQESSRSYK	121
Db	55	MSCGYAIGIGVNRFPYLCGRKNGAGFLIPFLTLIPAGVPLLELLESLQYISIGLG	114
Qy	122	VWISTSPMKGNGYVQAQAGCYILSYVYVIGLCYLTYLAMSFOATLPMAICQPEV-ENCY	179
Db	115	VWKLAPKFGVGLAAVALSPFLNITTYIYIISMAITLYLNSTTTLLPKQCDNPNTDRCF	174
Qy	180	PSDFPLLAAYNNITNGTSSAQVLYFLFTVLQSDGIEGGLGAPIVLYVLCFIFAMLVYGV	239
Db	175	SN-----YSMNNTNMTSAVVEFEFERNHQMOTDGLDKP-GQIRPPLAITLAIAMILVYFC	228
Qy	240	VARGVSSSKRAYFLALFPYVYVMTLTFTTILLGANGDILGFYTPQMAKLELGVWYSA	299
Db	229	IMKGVGMTGKVVYSATPYMTLLIFPRGTLPGAEGILFYTPNFKLSDSEWLDADA	288
Qy	300	VTQVFFSLTYVGTGPIIMESSYNGRHNHYRPMVITVTLDTPTSLSGCTIPGILGNLAYE	359
Db	289	ATQLEFYSIGLIGSLIALGYSNPNHNVTDSIIYCCINCSCTSMFAGVYSIVGPMHV	348
Qy	360	LNSEYGVYVAGGSIAPISIPDAIAKTFQQLSVLFFLMSVYLGIGSSVALLSTFTL	419
Db	349	TKRSIAD-VAASGGGLAFLEAVYQTLPISLPALLFEFSLMLMGIDISQCTVEGFTTA	407.
Qy	420	AMDAFPFY-----PYYVSAMTCSCGFLGLVYCTPGGQYILIELVDHYGTFEL-VLFCAIS	474
Db	408	LVDEYPRLLRRRRELFIAAV-CIISYLGISLNTQGGIYVPLKLDYISASGMSLLFLVEF	466
Qy	475	ELAGVFWIYTGLENLCIDIEPMGLKKTGAYWFLCWCVTIPATIMTVFFPALLASNNLVFGD	534
Db	467	ECVSIWMYGVNRRFYDNIQEWVGSPPCIWMLKCSFFPIIVAGVFIFSAQMPTLWG-	525
Qy	535	NYVYPTAGYVSGYVLMFLGMEFVPFGIGFISLYKRT--GFPESETIKKAFHS-----KPSW	587
Db	526	NYVFPKMGQGVGMALLSSMVLIP---GYMAYMFLALKGSLKQRIQVWVQPSDPTVREN	582
Qy	588	GP 589	
Db	583	GP 584	

  

RESULT 6			
Seq1_RAT	STANDARD:	PRT:	599 AA.
AC	P23978;		
DT	01-MAR-1992 (Rel. 21, Created)		
DT	01-MAR-1992 (Rel. 21, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Sodium- and chloride-dependent GABA transporter 1.		
GN	SLC6A1 OR GABT1 OR GAT-1.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
NCBI_Taxid	10116;		
RN	[1]		
RN	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
RC	TISSUE=Brain;		
RX	MEDLINE=90378307; PubMed=1975955;		
RA	Guastella J., Nelson N., Nelson H., Czyzyk L., Keynan S.,		
RA	Miedel M.C., Davidson N., Lester H.A., Kanter B.I.;		
RT	"Cloning and expression of a rat brain GABA transporter.";		
RL	Science 249:1303-1306(1990).		
RN	[2]		
RP	TOPOLOGY.		
RX	MEDLINE=97313439; PubMed=959433;		
RA	Clark J.A.;		
RT	"Analysis of the transmembrane topology and membrane assembly of the		
RL	GAT-1 gamma-aminobutyric acid transporter.";		
J. Biol. Chem.	272:14695-14704(1997).		
-1-	FUNCTION: TERMINATES THE ACTION OF GABA BY ITS HIGH AFFINITY		
CC	SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.		

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR  
 CC STIMULANTS SUCH AS AMPHETAMINES OR COCAINE.  
 CC -1- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER  
 CC FAMILY (SNF).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M59742; AAA63487.1; -  
 DR PIR: A35918; ACRGT.  
 DR InterPro: IPR000175; Na/ntran\_symport.  
 DR Pfam: PF00209; SNF. 1.  
 DR PRINTS: PR00176; NAMEUSMPORT.  
 DR PRODOM: PD000448; Na/ntran\_symport; 1.  
 DR PROSITE: PS00610; NA\_NEUROTRAN\_SYM\_1; 1.  
 DR PROSITE: PS00754; NA\_NEUROTRAN\_SYM\_2; 1.  
 DR PROSITE: PS50267; NA\_NEUROTRAN\_SYM\_3; 1.  
 KW Neurotransmitter transport; transport; Transmembrane; Glycoprotein;  
 KW Symport; Multigene family.  
 FT DOMAIN 1  
 FT TRANSMEM 53 73  
 FT DOMAIN 74 80  
 FT TRANSMEM 81 100  
 FT DOMAIN 101 123  
 FT TRANSMEM 124 144  
 FT DOMAIN 145 211  
 FT TRANSMEM 212 230  
 FT DOMAIN 231 256  
 FT TRANSMEM 257 291  
 FT DOMAIN 292 309  
 FT TRANSMEM 310 320  
 FT DOMAIN 321 342  
 FT TRANSMEM 343 374  
 FT DOMAIN 375 394  
 FT TRANSMEM 395 421  
 FT DOMAIN 422 440  
 FT TRANSMEM 441 456  
 FT DOMAIN 457 477  
 FT TRANSMEM 478 497  
 FT DOMAIN 498 517  
 FT TRANSMEM 518 535  
 FT DOMAIN 536 554  
 FT TRANSMEM 555 599  
 FT DOMAIN 176 176  
 FT CARBOHYD 181 181  
 FT CARBOHYD 184 184  
 SQ SEQUENCE 599 AA; 67001 MW; 4FEF85092DC1D045 CRC64;  
 Query Match 30.18; Score 1008; DB 1; Length 599;  
 Best Local Similarity 34.88; Pred. No. 3.1e-59;  
 Matches 211; Conservative 122; Mismatches 231; Indels 42; Gaps 13;  
 QY 3 DGVGVNGFESSEPEKMRSSQIS-LPPANNKALNDIDTDLAEPPERWMSNNTEEL 61  
 DB 10 DGGI-----STEVSEAPVASDKRKTLLVKKAGDLPDDTYMKGRDFL 54  
 QY 62 MSCIATSVGLGNWRPEPIAYQNGGAFIPYIVLVLLVKGPPYYLCEVGFSSRSRSYK 121  
 DB 55 MSCVGAIGLGNWRPFYILGKNGGAFILPFLITLIFAGVPLFLECSIGQYTSIGLG 114  
 QY 122 VMSISPRAMKGTGAQAAGCGYIISYYVITGLCLLYIAMSFOATLPAICQPEW--ENCY 179  
 DB 115 VMKLAFMFKGVGAAALVSEFNIYYIVIIISMAIYYIYNFTTLEPKQCDNPMNDRCF 174  
 QY 180 PSDPTLAASVNNITNGSSAQLEFLRIVYLDQSDIGSGLAGAPITWYLVLCFIAMLVWEGV 239

DB 175 SN-----YSLVNTNMTSAVVEFWRNMQHTDGLDKP-GQIRWPLATLAIAMVLYFC 228  
 QY 240 VARGVSSGKAAYALALEPPYVMVLTFTTITLIGADGILFEFTYPMQAKLEGWKSA 299  
 DB 229 IMKGWGMKGVYSAIYPIYMLITLIFRGYLGCAEGILFYITPFRKISDSEVMIDA 288  
 QY 300 VTQVFFSLVCTGPIINSSYNGFRHNIYRDAMIYVTLDPFTPSLSCCTIGILGNLAYE 359  
 DB 289 ATQLEFFSGIGLGLINAGSNHNMYNRISIIVCCINCSNCFAGVIFSIYGFMAHV 348  
 QY 360 LNSEVDYVGGGSLAFISIPDAIATPQPLSFVFLMMSVGLGSSVALLSTENTL 419  
 DB 349 TKRSIAD-VASGGLFLAYPEAVTQLPISPAIILFFSMLMLGIDSOQCTVEGPTA 407  
 QY 420 AMDAPRY-----PIVYVSAMTSCSGFLIGVYCPGGQYILELDVHGGFPL-VLPCAIS 474  
 DB 408 LVDEYPRILNRRLRELFIAAV-CIYSYLIGLSNITGGGYVFKLDYVSASMSLFLVVF 466  
 QY 475 ELAGVFIYGLLENCLDIEFMIGKRTGAYWRLCNGVITPAIMTYFFAYALLASNNLVFGD 534  
 DB 467 ECVISIMFYGVNRRYDNIQGEVSGRPMWMLCKSFPTPIVAGVFLFSANQMTPLMG- 525  
 QY 535 NYVPTAGYVSGYMLMLGMTFVPIGIFSLYKRT--GFSETIKAHFS----KPSW 587  
 DB 526 SYVEPKWGGQGVMLALSMVLIP---GYAMVFLTKSLKRLQVMIQPSEDIYVPEP 582  
 QY 588 GPRSPR 593  
 DB 583 GPEOPQ 588  
 RESULT 7  
 ID S644 MACMU STANDARD; PRT; 630 AA.  
 AC Q9MYXO;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Sodium-dependent serotonin transporter (5HT transporter) (5HTT).  
 GN SLC6A4.  
 OS Macaca mulatta (Rhesus macaque).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 CC Cercopithecinae; Macaca.  
 OX NCBI\_Taxid=9544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Midbrain;  
 RX MEDLINE=21126513; PubMed=11223167;  
 RA Miller G.M., Yalin S.M., De la Garza R. II, Goulet M., Madras B.K.;  
 RT "Cloning of dopamine, norepinephrine and serotonin transporters from  
 RT monkey brain: relevance to cocaine sensitivity."  
 RT Brain Res. Mol. Brain Res. 87:124-143(2001).  
 CC -1- FUNCTION: TERMINATES THE ACTION OF SEROTONINE BY ITS HIGH AFFINITY  
 CC SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR  
 CC STIMULANTS SUCH AS AMPHETAMINES OR COCAINE (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER  
 CC FAMILY (SNF).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF285761; AA97247.1; -  
 DR InterPro: IPR002437; 5HT\_transporter.  
 DR InterPro: IPR000175; Na/ntran\_symport.

DR Pfam; PF00209; SNF; 1.  
 DR Pfam; PF03491; SHT transporter; 1.  
 DR PRINTS; PR00176; NANUSMPORT.  
 DR ProDom; PD000448; Na/neutran\_sympt; 1.  
 DR PROSITE; PS00610; NA\_NEUROTAN\_SYM\_1; 1.  
 DR PROSITE; PS00754; NA\_NEUROTAN\_SYM\_2; 1.  
 DR PROSITE; PS50267; NA\_NEUROTAN\_SYM\_3; 1.  
 DR Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;  
 KW Symptom.  
 FT DOMAIN 1 87 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 88 108 1 (POTENTIAL).  
 FT TRANSMEM 116 135 2 (POTENTIAL).  
 FT TRANSMEM 160 180 3 (POTENTIAL).  
 FT DOMAIN 181 252 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 253 271 4 (POTENTIAL).  
 FT TRANSMEM 280 297 5 (POTENTIAL).  
 FT TRANSMEM 333 350 6 (POTENTIAL).  
 FT TRANSMEM 362 383 7 (POTENTIAL).  
 FT TRANSMEM 417 436 8 (POTENTIAL).  
 FT TRANSMEM 464 482 9 (POTENTIAL).  
 FT TRANSMEM 518 558 10 (POTENTIAL).  
 FT TRANSMEM 539 558 11 (POTENTIAL).  
 FT TRANSMEM 577 595 12 (POTENTIAL).  
 FT DOMAIN 596 630 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 630 AA; 70257 MW; 79DFB7E5E9EB6924 CRC64;

Query Match 29.8%; Score 998; DB 1; Length 630;  
 Best Local Similarity 37.0%; Pred. No. 1.5e-58;  
 Matches 226; Conservative 98; Mismatches 241; Indels 46; Gaps 15;

1 MNDGVNGVFSSSEKMEPRKSSQSLPPANNKALNDIDPLEAE--PPRAWVSNNT 58  
 38 VESGQISNGY-SAVSPGADDDTRHSIRAF-----TTVAELHOGREHWGKKV 86  
 59 EFLMSCIATSVGLGNWRRPPIAYONGGAGFLVPYVILLVGRPVYLLGVGGSSRN 118  
 87 DELLSVIGYAVDLGNWRRPPIAYONGGAGFLVPYVILLVGRPVYLLGVGGSSRN 146  
 119 SVKWN-SISPMKKGTVYQAAAGCCYLLSVYVIGLCILYLAWSQATLPAAICQPEW 176  
 147 CISTWRKICPIFKGIGVACIIATFYASTYNTIMAMALYLLISFTDQLPMTCKNSMT 206  
 177 -NCVPSDPTLAASVNNIT---NGTSSAQLYFLRYLQ--QSDGIGGLGAPITMYVLCLF 230  
 207 GNC-----TNFSEENITWTLHSTPAPAEFYRHHVQLHRSKGLQ-DLGISMQALACIM 260  
 231 IATLWVEGVAVAGVSKSKAYLALFPYVMTLFTTILLPGATDGLFFVTPQMAKL 290  
 261 LIFVIVPSIMGVKTSKVVWVATFPYIILSVLLVNGATLPGAMRGVLYFLKPMQKL 320  
 291 LELGWYSANVQVFFSLVVCGRPIIMFSSYNGFRINIRDAIVITLDTFTFLSGCTIF 350  
 321 LETGWIDAAQIFESLSPGFGVLLAFASYNKFNNCYODALVTSVNCMTSEVSGFYLF 380  
 351 GILGNLAELNSEVDVVGAGSTSLAFISYPDAIKTRQPOLFFSVLFLFMSVGLGSSV 410  
 381 TVLGTMALMRNEDVSEVAKDAGPSLFTFYAIAINMPASTFAIIFLMLTTLGLDSTF 440  
 411 ALLSTFNTLMDAPRPVPT-----VYMSAMTSCGGLGLVYCPGGGYILLEVHYDG 464  
 441 AGLEGVITAVLDEPFHIAKREMFVLANVITCFGSLVTLF---GGAYVKKLLEEVAT 497  
 465 TFLVFCALSELAGFWITVGLNCLDIEFMIGKRTGAVNRCKGVITPAIMTYFFVYL 524  
 498 GPVAVLVALEAVASWFEYGLTFQFCDVKEMLGFSFGWWRICWVAISPLFLITICSL 557  
 525 IASNNL-VFGDNVYVPTAGYVGYLMLFGMTFVIGIFESYKXRTGFSEETIKAFHS 583  
 558 MSPQLRLF--QYNPMSIILGYCIGISSFVCIFTYIAKRLIS-TPGIFKRRIIK----- 610  
 584 KPSWGRPSPRE 594

DB 611 --STPEPTE 619.  
 RESULT 8  
 ID S6A1\_MUSCO STANDARD; PRT; 598 AA.  
 AC P48057;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Sodium- and chloride-dependent GABA transporter 1.  
 GN SLC6A1 OR GABRI OR GAT-1.  
 OS Mus coxli (COCK's mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10098;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Retina;  
 RX MEDLINE=95049273; PubMed=7960586;  
 RA Ruiz M., Bgal H., Sathly V.P., Qian X.J., Sarkar H.K.;  
 RT "Cloning, expression, and localization of a mouse retinal gamma-  
 aminobutyric acid transporter."  
 RL Invest. Ophthalmol. Vis. Sci. 35:4039-4048(1994).  
 CC -1- FUNCTION: TERMINATES THE ACTION OF GABA BY ITS HIGH AFFINITY  
 CC SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- TISSUE SPECIFICITY: BRAIN.  
 CC -1- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR  
 CC STIMULANTS SUCH AS AMPHETAMINES OR COCAINE.  
 CC -1- SIMILARITY: BELONGS TO THE SODIUM:NEUROTANSMITTER SYMPORTER  
 CC FAMILY (SNF).  
 CC -----  
 CC This SNF-SYMPOT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U32178; AAA37663.1; -  
 CC MGD; MGI:95627; Gabt1.  
 DR InterPro; IPR00175; Na/neutran\_sympt.  
 DR Pfam; PF00209; SNF; 1.  
 DR PRINTS; PR00176; NANUSMPORT.  
 DR ProDom; PD000448; Na/neutran\_sympt; 1.  
 DR PROSITE; PS00610; NA\_NEUROTAN\_SYM\_1; 1.  
 DR PROSITE; PS00754; NA\_NEUROTAN\_SYM\_2; 1.  
 DR PROSITE; PS50267; NA\_NEUROTAN\_SYM\_3; 1.  
 KW Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;  
 KW Symptom; Multigene family.  
 FT DOMAIN 1 52 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 53 73 1 (POTENTIAL).  
 FT TRANSMEM 81 100 2 (POTENTIAL).  
 FT TRANSMEM 124 144 3 (POTENTIAL).  
 FT DOMAIN 145 211 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 212 229 4 (POTENTIAL).  
 FT TRANSMEM 238 255 5 (POTENTIAL).  
 FT TRANSMEM 291 308 6 (POTENTIAL).  
 FT TRANSMEM 320 341 7 (POTENTIAL).  
 FT TRANSMEM 374 393 8 (POTENTIAL).  
 FT TRANSMEM 421 439 9 (POTENTIAL).  
 FT TRANSMEM 456 476 10 (POTENTIAL).  
 FT TRANSMEM 497 516 11 (POTENTIAL).  
 FT TRANSMEM 535 553 12 (POTENTIAL).  
 FT DOMAIN 554 598 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 598 619 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 598 AA; 66783 MW; 78FEC99402EC71C3 CRC64;









Db	196	PMTSCNSMNTGNC-----TNFSEDNITWTLHSTSPAEFFTRHVLQHRSGLO-DLG	249
Qy	220	AIPIWIVLCLFIAMLMWEGVAVARGKSSGKAAYFLALFPYVWITLFTITIIILPGATDGI	279
Db	250	GLSMPLVICIMFIFIVTIVYFISIMKGVTSKVVWVATFPYIILLILVNGATILPGAMRGV	309
Qy	280	LEFVVPOMAKLLELGWYSAVTVQVPEPSLTVCGPIIMESSYNGFRNITRDAMITVTLDT	339
Db	310	LEYLFPNMOKLLETGWWDAAAOIFESLDPGFGVLLAFASYKNKFHNHCYODALVYSAVNC	369
Qy	340	FTSLSGCTIFGILGNLAYELNSSEGDVVGAGGTSIAFISYDPAIKTQPOLFSVLFL	399
Db	370	MSPFSGFIVTIVLGVMAEMRKEDYSEVAKADGPKSLFTIYEAIALNMASFFFAVETL	429
Qy	400	MMSVIGIGSVALLSTFNPLANDAPRPVTF-----VYNSAMTSCGCFLLGLVYCPGQ	453
Db	430	MITVIGLSTFEGVLGVTAVLDEPPHWAKRREWFVLGVITCFGSLVTLTF---GGA	486
Qy	454	YLLELVHDHGGFFLVLFCAISELAGFYVYIGHENCLDIEFMIGKRTGAYWELCWVTP	513
Db	487	YVVKLLEPAGPVALVIALIAVAVFWYVYINGOCSDYKEMKSPGFMFKICWVAISP	546
Qy	514	AIMTVFFRYALLASNNL-VFGDNNYYPRAGVSGYIMFLGMTFVPDIGISFLSKYKRTGT	572
Db	547	LFLPLIFCSFLMSPQLRFL--QYDYPFMSIILGYCIGHSFICIPITYITVYRLI-VTPT	603
Qy	573	FSETIKK 579	
Db	604	LKERLIK 610	

  

RESULT 12				
56A4_RAT				
ID	56A4_RAT	STANDARD:	PRT:	630 AA.
AC	P31652:	P23976:		
DT	01-JUL-1993 (Rel. 26, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Sodium-dependent serotonin transporter (5HT transporter) (5HTT).			
GN	SLC6A4.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
KN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=WiStar;			
RX	MEDLINE=92111740; PubMed=1765155;			
RA	Mayer W., Betz H., Schloss P.;			
RT	"Isolation of cDNAs encoding a novel member of the neurotransmitter			
RT	transporter gene family.";			
RL	FEBS Lett. 295:203-206(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain stem;			
RX	MEDLINE=92049754; PubMed=1944572;			
RA	Blakely R.D., Berson H.E., Fremieu R.T. Jr., Caron M.G., Peek M.M.,			
RA	Prince H.K., Bardley C.C.;			
RT	"Cloning and expression of a functional serotonin transporter from			
RT	rat brain.";			
RL	Nature 354:66-70(1991).			
RN	[3]			
RP	REVISIONS.			
RC	TISSUE=Brain stem;			
RA	Blakely R.D.;			
RP	Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92054541; PubMed=1948036;			
RA	Hoffman B.J., Mezey E., Brownstein M.J.;			
RT	"Cloning of a serotonin transporter affected by antidepressants.";			
RL	Science 254:579-580(1991).			
RN	[5]			

DB 151 WRKICPFGKIGVACIARIAYNTIIMALYLISSLDRLPMTSTGNSMNTGNC- 209  
 QY 180 PSDPTLAASVNNIT---NGTSSAOVLFRTVLO--QSDGIEGGAGIAPWVLCIFIAWL 234  
 DB 210 ---TNFADQNTITWHTSHPAEFELRLVLOHOSKQIG-DLGITSMOLITCIVLIFT 264  
 QY 235 WVEGVAVRGVYSSKKAAYFLALFFPVVMTLFTITLLPCATDGLFFVTPQAKLLELG 234  
 DB 265 VVYFSIMKGVKTSKVVWVATPFYIVLVSILVAGALLPGAMRGVVFYELKPNMOKLLETG 324  
 QY 295 VVYSAVVOVFESLTVCTGPIIMSSYNGEFRNHYRITDPTDFTSPFGSCGIFGILG 354  
 DB 325 VVMDAAOIFESLSPGCGVLAFASYNKFNKNCQDLVYSVNCMTSPFSGPIFYVLG 384  
 QY 355 NLAYELNSEVDVYAGAGTSLAFISYDPAIAKTFQPLQFVSLFELMSVGLIGSSVALLS 414  
 DB 385 YMAEMRNEDSEVAKDAGPSLFTTVAEAIANMPASFEFALIFELMLITGLDSTFAGLE 444  
 QY 415 TFFNLAMDAPRVP-----VYMSAMTSCGFLGLVYCPGQYILLEVHDHGGFFLV 468  
 DB 445 GVITAVLDEPFIHAKRREMFVLIVITCVLGSLLT---TSGAVVVTLLSEYATGPV 501  
 QY 469 LECAISELAGVFWIYGLENLCLDIEFMLKRTGAYWRLCMGVITPAITVFFYALLASN 528  
 DB 502 LTVALLIEAVSWFYGTIFQFCSVDYKEMLGFSGFWRLCWALISPLFLIFICSFLMSPP 561  
 QY 529 NL-VFGDNVYVPTAGVYSGYMLFLGMEFVPIGIGFSLYKRYGTFFSETTKAHKSPW 587  
 DB 562 QLRLT---QYNYPHMSIVLYGICIGMSVYICPTIYIYRLIS--TPGLKERIK-----SI 612  
 QY 588 GPRSPRE 594  
 DB 613 TPEPTE 619

## RESULT 13

SEAC\_RAT STANDARD: PROT: 614 AA.

AC P48056: 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Sodium- and chloride-dependent betaine transporter (Na+/Cl-  
 DE betaine/GABA transporter).  
 GN SLC6A12.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=Sprague-Dawley; TISSUE=Liver;  
 RX MEDLINE=97019277; PubMed=8665807;  
 RA Burnham C.E., Buerk B., Schmidt C., Bucuvalas J.C.;  
 RT "A liver-specific isoform of the betaine/GABA transporter in the rat:  
 RT cDNA sequence and organ distribution."  
 RT Blochim. Biophys. Acta 1284:4-8(1996).  
 CC - FUNCTION: TRANSPORTS BETANE AND GABA. MAY HAVE A ROLE IN  
 CC REGULATION OF GABAERGIC TRANSMISSION IN THE BRAIN THROUGH THE  
 CC REUPTAKE OF GABA INTO PRESYNAPTIC TERMINALS, AS WELL AS IN OSMOTIC  
 CC REGULATION (BY SIMILARITY).  
 CC - SUBCELLULAR LOCATION: Integral membrane protein.  
 CC - SIMILARITY: BELONGS TO THE SODIUM-NEUROTRANSMITTER SYMPORTER  
 CC FAMILY (SNF).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: U028927; AAC52867.1; ALT\_INIT.  
 DR InterPro: IPR000175; Na/alan\_symport.  
 DR Pfam: PF00209; SNF\_1.  
 DR PRINTS: PR00176; NANEUSMPOR.  
 DR ProDom: PD000448; Na/alan\_symport; 1.  
 DR PROSITE: PS00610; NA\_NEUROTRAN\_SYM\_1; 1.  
 DR PROSITE: PS00754; NA\_NEUROTRAN\_SYM\_2; 1.  
 DR PROSITE: PS0267; NA\_NEUROTRAN\_SYM\_3; 1.  
 DR Neurotransmitter transport; Transport; Glycoprotein;  
 KW Symport.  
 FT DOMAIN 1  
 FT TRANSMEM 45 44  
 FT TRANSMEM 65 65  
 FT TRANSMEM 73 92  
 FT TRANSMEM 117 137  
 FT TRANSMEM 138 210  
 FT TRANSMEM 211 229  
 FT TRANSMEM 238 255  
 FT TRANSMEM 291 308  
 FT TRANSMEM 320 341  
 FT TRANSMEM 374 393  
 FT TRANSMEM 423 441  
 FT TRANSMEM 458 478  
 FT TRANSMEM 499 518  
 FT TRANSMEM 538 556  
 FT DOMAIN 557 614  
 FT CARBOHYD 171 171  
 FT CARBOHYD 183 183  
 SQ SEQUENCE 614 AA; 69748 MW; 14C78DE5E1D808B CRC64;  
 Query Match 29.28; Score 979; DB 1; Length 614;  
 Best Local Similarity 36.18; Pred. No. 2.ee-57;  
 Matches 216; Conservative 111; Mismatches 233; Indels 38; Gaps 13;  
 QY 39 IDDTDLAEPPERPERVWVSNNTIEFLMSCATSYSGNFWRPFIYONGGAFVLYVLL 98  
 DB 25 MDQKDKD-QVMDRQMTNKMFEVLSVAGELITLGNWRFYLYCKNGGAFPIFYFFF 83  
 DB 99 LVGRPVYLLCVALGQESSRNSVKW--SISPMKGTGYAQAAGCYILSYVVICGLCY 157  
 DB 84 SCGIFVFLFVALGQYSSQGSVTAAMRKICPLDGLGASVVISYLLIYIILAMVLF 143  
 QY 158 LAMFQATLFWAICQPEW--ENCYP---SDPTLAASVNNITNGTSSQQLFLRTVLOSD 212  
 DB 144 LFSFTWELPWTCTNSMTEHCVDPLNSTRASYS--NFTSPVMEWERVIGITS 201  
 QY 213 GIEGGLAPWYLYVLCFIAMLVGAVRGVYSSKKAAYFLALFFVYVMTLFTITIL 272  
 DB 202 GIH-DLSLRKRELALCLILAMITCYFCIMKGVKSGVYFTAFPIMLIILIRVTL 260  
 QY 273 PGATDGLIFEVTPQMAKLELGVYSAVTOVFSLVTCGPIIMSSYNGEFRNHYRITDAM 332  
 DB 261 PGAVOGIYFLKPLDLRLKDPQVWMDAGTOIFFSFALCQGLTFLAGSYNNYHNHCYRDI 320  
 QY 333 IVTLDFITSLSCGTFGLIGNLAYELNSVGVGVAGGTSALAFISYDPAIAKTFQPL 392  
 DB 321 ALCELNSTSEVAFVFSILGFMAOEGVPISE--VABSGGLFIAPRAVYTMPPLSOL 379  
 QY 393 FSVLEFLMYSVGLIGSSVALLSTFNTLAMDAP-----RVPTVYMSAMTSCGFLIG 444  
 DB 380 WSCFLFLMLPLGLDSDQVCMELCVTASMDMFQPOLKRRGRELILALVAIVC--YLMG 436  
 QY 445 LVYCTPGQYILLEVDHYGCT-FLVLCFCAISELAGVFWIYGLENLCLDIEFLMKRTGAY 503  
 DB 437 LLVTEGMYIFQFLDYVASSGICLLFSLFEVICYMGVADRFYDVEDMIDYRPWPL 496  
 QY 504 WRLCMGVITPAIMTVFFYALLASNNIVEGNNVYVPTAGVYSGYMLFLGMEFVPIGIGF 563  
 DB 497 VKISWLFITLPOLCLATEFFSLSKYTPLYNNVYTPSGSITGLWLFSSMACYPLPIII 556  
 QY 564 SLVYRTGTFFETIKKAHFKSPW---GPRSPRREMMQKABAKALROKMNTSRVYK 618  
 DB 557 TLAKTQ-GSFKKRRLRLTPPSLPQPERSPD-----GSAQNCSTSPVK 602

RESULT 14  
S6A1 MOUSE STANDARD; PRT; 598 AA.

AC P31648;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Sodium- and chloride-dependent GABA transporter 1.  
GN SLC6A1 OR GABT1 OR GAT-1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92335351; PubMed=1631167;  
RA Liu Q.-R., Mandliyan S., Nelson H., Nelson N.;  
RL "A family of genes encoding neurotransmitter transporters.";  
Proc. Natl. Acad. Sci. U.S.A. 89:6639-6643(1992).  
CC -1- FUNCTION: TERMINATES THE ACTION OF GABA BY ITS HIGH AFFINITY  
SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- TISSUE SPECIFICITY: BRAIN.  
CC -1- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR  
STIMULANTS SUCH AS AMPHETAMINES OR COCAINE.  
CC -1- SIMILARITY: BELONGS TO THE SODIUM-NEUROTRANSMITTER SYMPORTER  
FAMILY (SNP).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sdb.cn/announce/>  
or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
-----  
DR EMBL: M92377; -; NOT\_ANNOTATED\_CDS.  
DR EMBL: M92378; -; NOT\_ANNOTATED\_CDS.  
DR PIR: F46027; F46027.  
DR MGD: MGI:95627; Gatl1.  
DR InterPro: IPR000175; Na/ntran\_symport.  
DR Pfam: PF00209; SNF; 1.  
DR PRINTS: PRO0176; NANEUSMPORT.  
DR ProDom: PD000448; Na/ntran\_symport; 1.  
DR PROSITE: PS00610; NA-NEUROTRAN\_SYMPT\_1; 1.  
DR PROSITE: PS00754; NA-NEUROTRAN\_SYMPT\_2; 1.  
DR PROSITE: PS50267; NA-NEUROTRAN\_SYMPT\_3; 1.  
KW Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;  
KW Symport; Multigene family.  
FT DOMAIN 1 52 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 53 73 1 (POTENTIAL).  
FT TRANSMEM 81 100 2 (POTENTIAL).  
FT TRANSMEM 124 144 3 (POTENTIAL).  
FT DOMAIN 145 211 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 212 229 4 (POTENTIAL).  
FT TRANSMEM 238 255 5 (POTENTIAL).  
FT TRANSMEM 291 308 6 (POTENTIAL).  
FT TRANSMEM 320 341 7 (POTENTIAL).  
FT TRANSMEM 374 393 8 (POTENTIAL).  
FT TRANSMEM 421 439 9 (POTENTIAL).  
FT TRANSMEM 456 476 10 (POTENTIAL).  
FT TRANSMEM 497 516 11 (POTENTIAL).  
FT TRANSMEM 535 553 12 (POTENTIAL).  
FT DOMAIN 554 598 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 176 176 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CARBOHYD 181 181 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CARBOHYD 184 184 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 598 AA; 66841 MW; 893C3575DB99BD00 CRC64;

Query Match 29.2%; Score 978.5; DB 1; Length 598;  
Best Local Similarity 34.3%; Pred. No. 2.7e-57;

Matches 208; Conservative 123; Mismatches 232; Indels 43; Gaps 13;

QY 3 DQVNGGFESSEPKKRRSQIS--LPPANRAALNDIDDTLEAPPERAWNNIIEPL 61  
DDB 10 DQOI-----STEVSAPVASDKPKTLVYVYOKKADLDLPRODTWKGRFDL 54  
QY 62 MSCIAITVGLGVNWRPFFLAIONGGAFLLPVYVILLGVKVVYLLGVGGFSSRNRYK 121  
DDB 55 MSCVGAIGLVNWRPFLCGKNGGAFLLPFLTLIFAGVPLFLFECISLGQYTSIGIG 114  
QY 122 WMSISPMKGTGYAAGCGYLLSYVYVGLCLVYLLMSFPQATLPMHICQPEW--ENCV 179  
DDB 115 VNNMAMPKGVAAVAALVSLFNLITYIYISAIYLLNSFTTLLPMQCDNPMWTDKCF 174  
QY 180 PEDPFLAASVNNITNGTSSAQLFLRTVYLOQSDGIEGGIAPIMVLYLCLFAMLVFV 239  
DDB 175 SN-----YSLVNTTMTSTAVVEFWERNHQMOTGDH--KPGIICRLATLTAIVLYFC 227  
QY 240 VARGVSSGKAAYFLALPPYVMTLFTTTTLTLCATDGIILFFVPPQAKLLEGVWSA 299  
DDB 228 IWKGVGTGKVVYFSAITPYIMLTILFRGVTLPAKKGILFYITPNFRKLSDSSEVIFA 287  
QY 300 VTQVEFSLVCTGPIIMSSYNGFRNHYRDAMIYTTLDFTSFLSGCTIFGILGNAYE 359  
DDB 288 ATQIFESYGLGLSLALGSVNSFNHNYRDSIIVCCINSGTSMFAGVIFSIVGMAYH 347  
QY 360 INSEVDVYAGGTSIAFTSDAIAKTFQPLESVLFLMMSVIGIGSSVALLSTFNTL 419  
DDB 348 TKRSIAD--VAASGPILAFIAPVAVTOLPISPLAILFESMLMIGDISQCTVGEFYTA 406  
QY 420 AMDAPRY-----PYVMSAMTSCGFLGLVYCPGSGGYITIELVDHVGITL-VIFCAIS 474  
DDB 407 LVDEPRLLRNRELFIANV-CIYSYLLGSLNITGGVYTRLPDYISAGSLSLELVF 465  
QY 475 ELAIVFYIGLEMLCLDIEFMIGKRTGAYWRLCWGVTPAINTVFYFALLASNNLVYGD 534  
DDB 466 ECVSISWYGVNRFYNDIOEMVSRPCIMMKLCSFFPIIVAGVLEFSAVQMPPLTMG 524  
QY 535 NVVYPTAGVSGYLMFLFGMTFVPIGIGFSLXKYT--GTSETIKKAFHS-----KRW 587  
DDB 525 SYVPPKMGQGVWMLALSSMLVLP--GYMAYMFLTGLSKORLVMIQPSIEDIVRPN 581  
QY 588 GPRSPR 593  
DDB 582 GPEQPO 587

RESULT 15  
S6AB\_HUMAN STANDARD; PRT; 632 AA.

AC P48066;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 16-FEB-1996 (Rel. 33, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Sodium- and chloride-dependent GABA transporter 3.  
GN SLC6A1 OR GABT3 OR GAT3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Fetal brain;  
RX MEDLINE=95179472; PubMed=7874447;  
RA Borden L.A., Murall Dhar T.G., Smith K.E., Branchek T.A.,  
Gluchowski C., Welshank R.L.;  
RT Cloning of the human homologue of the GABA transporter GAT-3 and  
RT identification of a novel inhibitor with selectivity for this site.;  
RL Recept. Channels 2:207-213(1994).  
CC -1- FUNCTION: TERMINATES THE ACTION OF GABA BY ITS HIGH AFFINITY  
SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- TISSUE SPECIFICITY: WIDESPREAD DISTRIBUTION IN THE BRAIN.



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2003, 09:24:45 ; Search time 87 Seconds  
(without alignments)  
1499.170 Million cell updates/sec

Title: US-09-991-458-2  
Perfect score: 3354  
Sequence: 1 MNDGVNGGFESSEPKMEPR.....TSRVKHLWYITGAYRRIN 633

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: SPTREMBL\_21.\*
- 2: sp.archaea.\*
- 3: sp.bacteria.\*
- 4: sp.fungi.\*
- 5: sp.human.\*
- 6: sp.invertebrate.\*
- 7: sp.mammal.\*
- 8: sp.mhc.\*
- 9: sp.organelle.\*
- 10: sp.phage.\*
- 11: sp.plant.\*
- 12: sp.todent.\*
- 13: sp.virus.\*
- 14: sp.vertibrate.\*
- 15: sp.unclassified.\*
- 16: sp.virus.\*
- 17: sp.bacteriap.\*
- 17: sp.archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3354	100.0	633	5	09U5A9
2	3084	91.9	634	5	076188
3	1533	45.4	593	5	09W4C5
4	1510.5	45.0	654	5	095T51
5	1492	44.5	651	5	09V5B3
6	1472.5	43.9	639	5	09VJ4
7	1461.5	43.6	739	5	09NKA8
8	1414	42.2	629	5	09V690
9	1389.5	41.4	593	5	09VSV2
10	1238.5	36.9	615	5	08T7E1
11	1089.5	32.5	791	11	091ZQ2
12	1067.5	31.8	415	5	08SYT9
13	1030	30.7	638	11	09JMA9
14	1030	30.7	674	5	08WPM9
15	1027	30.6	638	11	09D317
16	1025	30.6	638	11	091Y60

17	1024	30.5	638	11	09R183	09R183 mus musculus
18	1010	30.1	642	4	09UN76	09UN76 homo sapien
19	1003.5	29.9	598	13	073771	073771 raja sp. 9a
20	999	29.8	598	13	091494	091494 torpedo cal
21	997.5	29.7	631	6	09GMA5	09GMA5 ovis aries
22	979	29.2	611	13	091502	091502 torpedo mar
23	971	29.0	581	5	023969	023969 drosophila
24	961	28.7	614	11	08VCS9	08VCS9 mus musculus
25	961	28.7	630	13	042482	042482 rana catesb
26	960	28.6	640	11	08VBM1	08VBM1 mus musculus
27	956	28.5	602	4	08TCC2	08TCC2 homo sapien
28	956	28.5	625	13	091BH6	091BH6 cyprinus ca
29	954.5	28.5	622	13	091503	091503 torpedo mar
30	950	28.3	629	13	09DE16	09DE16 oreochromis
31	949.5	28.3	621	11	091W12	091W12 mus musculus
32	947.5	28.2	629	13	090ZY1	090ZY1 brachydanio
33	947.5	28.2	1201	5	09W10	09W10 drosophila
34	945	28.2	633	11	08VC47	08VC47 mus musculus
35	944	28.1	635	13	09DGN5	09DGN5 gallus gall
36	942.5	28.1	620	6	09GJ76	09GJ76 macaca fasc
37	939	28.0	619	11	09J341	09J341 mus musculus
38	938.5	28.0	620	6	09GJ75	09GJ75 salmuri sci
39	938	28.0	617	11	063380	063380 rattus norv
40	936.5	27.9	620	6	09GJ77	09GJ77 salmuri sci
41	934.5	27.9	597	11	09WTR4	09WTR4 rattus norv
42	934	27.8	619	11	09R0X6	09R0X6 mus musculus
43	931.5	27.8	617	11	08R212	08R212 mus musculus
44	929.5	27.7	567	11	09WTR3	09WTR3 rattus norv
45	929.5	27.7	597	5	025512	025512 manduca sex

## ALIGNMENTS

RESULT 1  
ID 09U5A9 PRELIMINARY: PRT; 633 AA.  
AC 09U5A9;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Amino acid transporter/amino acid-gated channel for sodium/potassium ions.  
GN CAATCH1.  
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; OC Sphingioidea; Sphingidae; Sphinginae; Manduca.  
OX NCBI\_TaxID=7130;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=WIDGUT;  
RX MEDLINE=20387341; PubMed=10829035;  
RA Feldman D.H., Harvey W.R., Stevens B.R.;  
RT "A novel electrogenic amino acid transporter is activated by K+ or Na+, is alkaline pH-dependent, and is Cl--independent.";  
RL J. Biol. Chem. 275:24518-24526(2000).  
DR EMBL: AF013963; AAF18560.1; -;  
DR InterPro: IPR000175; Na/ntan\_symport.  
DR Pfam: PF00209; SNE; 1.  
DR PRINTS: PR00176; NANEUSMPORT.  
DR PRODOM: PDD00448; Na/ntan\_symport; 1.  
DR PROSITE: PS50267; NA\_NEUTRAN\_SYMPT\_3; 1.  
SQ SEQUENCE 633 AA; 69934 MW; 597EDC58E8FDFD CRC64;

Query Match 100.0%; Score 3354; DB 5; Length 633;  
Best local Similarity 100.0%; Pred. No. 1.4e-259;  
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNDGVNGGFESSEPKRSSQISLPANKKALNDIDDTDLAEPPERWVMSNIEF 60  
|||||  
DB 1 MNDGVNGGFESSEPKMEPKRSSQISLPANKKALNDIDDTDLAEPPERWVMSNIEF 60  
|||||

```

QY 61 LMSCIATSVGLGNWRPFPIAYONGGAFILVYVILLVKGPRYYLECVLGFQFSRNSV 120
DB 61 LMSCIATSVGLGNWRPFPIAYONGGAFILVYVILLVKGPRYYLECVLGFQFSRNSV 120
QY 121 KWSISPMKGTGVAQAAGCGIILSYVVICGLLYLAMSFOATLPMAICQPEMENCVP 180
DB 121 KWSISPMKGTGVAQAAGCGIILSYVVICGLLYLAMSFOATLPMAICQPEMENCVP 180
QY 181 SDPTLAASVNNITNGTSSAQLYFLRTVYLOQSDGIEGGIGAPIMYLVLCLFLAMLMVGVY 240
DB 181 SDPTLAASVNNITNGTSSAQLYFLRTVYLOQSDGIEGGIGAPIMYLVLCLFLAMLMVGVY 240
QY 241 ARGVSSGKAAYFLALFPYVMTLFTITILLPGATDGLFEFVTPOMAKLLELGVWYSAY 300
DB 241 ARGVSSGKAAYFLALFPYVMTLFTITILLPGATDGLFEFVTPOMAKLLELGVWYSAY 300
QY 301 TOVEFSLVCTGPTIMSSYNGFRNRYRDAMIYTTLDFTSFLSGCTIGILGNLAYEL 360
DB 301 TOVEFSLVCTGPTIMSSYNGFRNRYRDAMIYTTLDFTSFLSGCTIGILGNLAYEL 360
QY 361 NSEYGDVVGAGTSLAFISYPDALAKTFQPOLFSVLEFLMSSVIGISSVALLSTENTLA 420
DB 361 NSEYGDVVGAGTSLAFISYPDALAKTFQPOLFSVLEFLMSSVIGISSVALLSTENTLA 420
QY 421 MDAPRPVPTVYMSAMTSCGFLGLVYCTPGGQYILELVHYGTFVLVCAISELAGVF 480
DB 421 MDAPRPVPTVYMSAMTSCGFLGLVYCTPGGQYILELVHYGTFVLVCAISELAGVF 480
QY 481 WIYLENLCDIEFMLGKKTGAYWRLCMGVTPPAIMTVFEYALLASNNLVFGNNVYPT 540
DB 481 WIYLENLCDIEFMLGKKTGAYWRLCMGVTPPAIMTVFEYALLASNNLVFGNNVYPT 540
QY 541 AGVYGYLMLFLGTMFVPIGIFSLYKRYRTGTFSETIKKAFHSPKSPGPRSPRRERMQ 600
DB 541 AGVYGYLMLFLGTMFVPIGIFSLYKRYRTGTFSETIKKAFHSPKSPGPRSPRRERMQ 600
QY 601 FKAFAKALROKMNTSRVKHLMYSTTGAYRRNIN 633
DB 601 FKAFAKALROKMNTSRVKHLMYSTTGAYRRNIN 633

RESULT 2
076188 PRELIMINARY; PRT: 634 AA.
AC 076188;
DB 01-NOV-1998 (TREMBlrel. 08, Created)
DB 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DB 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DB Potassilum coupled amino acid transporter.
DB KATL.
DB Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
DB Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
DB Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
DB Sphingioidea; Sphingidae; Sphinginae; Manduca.
DB NCBI_TaxId=7130;
DB [1]
DB SEQUENCE FROM N.A.
DB MEDLINE=98226826; PubMed=9560287;
DB Castagna M., Shayakul C., Trotti D., Sacchi V.F., Harvey W.R.,
DB Hediger M.A.;
DB "Cloning and characterization of a potassilum-coupled amino acid
DB transporter.";
DB Proc. Natl. Acad. Sci. U.S.A. 95:5395-5400(1998).
DB EMBL: AF006063; AAC24190.1;
DB InterPro: IPR000175; Na/nitran-symport.
DB Pfam: PF00209; SNF. 1.
DB PRINTS: PRO0176; NANEUSMPORT.
DB ProDom: PD000448; Na/nitran-symport. 1.
DB PROSITE: PS0267; NA_NEUTROTAN_SYM_3; 1.
DB SEQUENCE 634 AA; 70239 MW; 28EA7A027D240654 CRC64;

```

```

Matches 571: Conservative 31: Mismatches 31: Indels 0: Gaps 0:
QY 1 MNDQVNGGFESSSEPKMKPKRSSQISLEPPANNKALNDIDTDLAEPPERWVNSNIEF 60
DB 1 MNDQVNGGFESSSEPKMKPKRSSQISLEPPANNKALNDIDTDLAEPPERWVNSNIEF 60
QY 61 LMSCIATSVGLGNWRPFPIAYONGGAFILVYVILLVKGPRYYLECVLGFQFSRNSV 120
DB 61 LMSCIATSVGLGNWRPFPIAYONGGAFILVYVILLVKGPRYYLECVLGFQFSRNSV 120
QY 121 KWSISPMKGTGVAQAAGCGIILSYVVICGLLYLAMSFOATLPMAICQPEMENCVP 180
DB 121 KWSISPMKGTGVAQAAGCGIILSYVVICGLLYLAMSFOATLPMAICQPEMENCVP 180
QY 181 SDPTLAASVNNITNGTSSAQLYFLRTVYLOQSDGIEGGIGAPIMYLVLCLFLAMLMVGVY 240
DB 181 SDPTLAASVNNITNGTSSAQLYFLRTVYLOQSDGIEGGIGAPIMYLVLCLFLAMLMVGVY 240
QY 241 ARGVSSGKAAYFLALFPYVMTLFTITILLPGATDGLFEFVTPOMAKLLELGVWYSAY 300
DB 241 ARGVSSGKAAYFLALFPYVMTLFTITILLPGATDGLFEFVTPOMAKLLELGVWYSAY 300
QY 301 TOVEFSLVCTGPTIMSSYNGFRNRYRDAMIYTTLDFTSFLSGCTIGILGNLAYEL 360
DB 301 TOVEFSLVCTGPTIMSSYNGFRNRYRDAMIYTTLDFTSFLSGCTIGILGNLAYEL 360
QY 361 NSEYGDVVGAGTSLAFISYPDALAKTFQPOLFSVLEFLMSSVIGISSVALLSTENTLA 420
DB 361 NSEYGDVVGAGTSLAFISYPDALAKTFQPOLFSVLEFLMSSVIGISSVALLSTENTLA 420
QY 421 MDAPRPVPTVYMSAMTSCGFLGLVYCTPGGQYILELVHYGTFVLVCAISELAGVF 480
DB 421 MDAPRPVPTVYMSAMTSCGFLGLVYCTPGGQYILELVHYGTFVLVCAISELAGVF 480
QY 481 WIYLENLCDIEFMLGKKTGAYWRLCMGVTPPAIMTVFEYALLASNNLVFGNNVYPT 540
DB 481 WIYLENLCDIEFMLGKKTGAYWRLCMGVTPPAIMTVFEYALLASNNLVFGNNVYPT 540
QY 541 AGVYGYLMLFLGTMFVPIGIFSLYKRYRTGTFSETIKKAFHSPKSPGPRSPRRERMQ 600
DB 541 AGVYGYLMLFLGTMFVPIGIFSLYKRYRTGTFSETIKKAFHSPKSPGPRSPRRERMQ 600
QY 601 FKAFAKALROKMNTSRVKHLMYSTTGAYRRNIN 633
DB 601 FKAFAKALROKMNTSRVKHLMYSTTGAYRRNIN 633

RESULT 3
09W4C5 PRELIMINARY; PRT: 593 AA.
AC 09W4C5;
DB 01-MAY-2000 (TREMBlrel. 13, Created)
DB 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DB 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DB CG3252 protein.
DB Drosophila melanogaster (Fruit fly).
DB Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
DB Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
DB Ephyroidae; Drosophilidae; Drosophila.
DB NCBI_TaxId=7227;
DB [1]
DB SEQUENCE FROM N.A.
DB STRAIN=BERKELEY;
DB MEDLINE=20196006; PubMed=10731132;
DB Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
DB Mananickes P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
DB George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
DB Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
DB Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
DB Wen K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
DB Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
DB Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

```

[illegible]

Oy		456	LELVDRYGGTFLVLCFAISGLAGVTWITGLENLCLDIDIEMLGKTKGAIVRLCMGYTPAI	515
Dd		417	LTLVDFFGGTYVFILIAIFELAGIWWVYLQNFCDIDEIMCRNRSLTWRCVCSFTPYM	476
Oy		516	MTVTFPVALLASNNLVFGDNVYVPYTAGVSGYLMFLMGTFWPDIGESLYKYRTGTSE	575
Dd		477	MIIIFYSMWTIEPIKYSLEY-PPELANIACGMILFALIGAOPPLMWISRHPOGYWK	535
Oy		576	TIKAFNSKPSWGSPRPRFREREMOKRAEAKKALR-QKNMTSYVKHIMWSI	624
Dd		536	SIKASLKPDSRMGWGPANPEIRREWRVFNKOKAARATQKDTSKLGFWRKV	585
 RESULT 4 O95T51 PRELIMINARY; PRT: 654 AA.				
ID	O95T51			
AC	O95T51:			
Dt	01-DEC-2001 (TREMBLrel. 19, Created)			
Dt	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
Dt	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
Dt	G21837P.			
DE	GN	CG1698.		
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
CC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
NCBI_TaxId=7227;	Ephydroidea; Drosophilidae; Drosophila.			
RN	(1)			
RC	SEQUENCE FROM N.A.			
RC	STRAIN-BEREKEY,			
RA	Stapleton M., Broksstein P., Hong L., Aghayani A., Carlson J.,			
RA	Champe M., Chavez C., Dorsett V., Farfan D., Flise E., George R.,			
RA	Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,			
RA	Nunoo J., Pacleib J., Paragas V., Park S., Phouenavong S., Wan K.,			
RL	Yu C., Lewis S.E., Rubin G.M., Gelinker S.;			
RL	Submitted (Oct-2001) to the EMBL/Genbank/DDBJ databases.			
DR	EMBL: AY060329; AAL25368.1; -.			
DR	FlyBase: FBgn0033443; CG1698.			
DR	InterPro: IPR000175; Na/_ntran_symport.			
DR	Pfam: PF00209; SNF; 1.			
DR	Prodomb: PD000448; Na/_ntran_symport; 1.			
DR	PROSITE: PS00267; NA_NEUROTRAN_SYM_3; 1.			
SQ	SEQUENCE 654 AA; 72438 MW; 4385D76627C68758 CRC64;			
 Query Match 45.0%; Score 1510.5; DB: 5; Length 654; Best Local Similarity 48.8%; Pred. No. 2.6e-112; Matches 309; Conservative 97; Mismatches 186; Indels 41; Gaps 13;				
Oy		11	ESSEPKKEPRRSOISLPAPNNAKALNDIDPDLEAEP--ERMWNSNIIEFMSCATS	68
Dd		38	ESSNGKKTTLSGSGSAPPBGHS-----PSCVEAGQGCKRDSNNNDIEFLMCIALS	91
Oy		69	VGLGNWRFPEFIAYONGGAFVLPPYIVILLVGKPVYYLECVIYGORSSRNYSKVMSIPA	128
Dd		92	VGLGNWRFPEFTLNLNGGAFVIPYLIVLLGVKPYYLEMGLGFOPSBSGSVXVDESPI	151
Oy		129	MKGTVGYQAAGCGIILSYVVVICGLCYLAMSFOATIPWAICOPWM-ENCYPSPDTLAA	187
Dd		152	MKGIGYQAVLATGIVTTVYATLALTRIVYDSFYPLPMSTYCREMGTECDLSGQEAS	211
Oy		188	SVNNITNG---TSSAOQLFRTVYLOQSDGIEGLADIVYLVCFLPAMLVNGVVARCV	244
Dd		212	RATSLAGSGVNTTSSEFTNIIILREKASIDDGIGYPSMSIALALAAVMIVAGIMEKW	271
Oy		245	KSSGKAAYFLAPRYVMITLFITTIILPGATDGLFEFPVPMNAKLLELCAVMSAVTOYE	304
Dd		272	KSSGKAAYFLAPRYVMIVLVLRALILPGAFDGVILFKRPQHKKLLEPQVWAAVTOVE	331
Oy		305	FSLPYCPTPIIMFSSYNGFRHNITYRDAMIYTTLDFTSPISGCTIFSIIGLAYELN-SE	363
Dd		332	FSALACPONIIMAYSNRGHNITYRDANIYTTLDFTFSLISGVIITFIIGLAYENNTTD	391
Oy		364	VGDVVGAGTSLAISTYDALAK-TFOQOLFVSVELFMMSVILGISGSAALLSTENTLAM	422

```

Db      392  IASVNV-GGPGIAFISYPDAIAKFKWLPQLEVSLEFLEMLVGLGSSNGMASCSTYVKD
Qy      423  APRRPVTVMSAMTSCGFLGLVYCTPGGOYILEVDHYGTFLVFCALSGLAGVMI
Db      451  QGCHLKNMTVVGVIAVGFGLLTYTPGGOFLLLVDFGVFVALVALAIFELVLTAMI
Qy      483  YLENLCDIDIEFMKKTGAWYRLCGVITPAIMTVTFVYALLASNNLVFGDNVYPTAG
Db      511  YGVKRLCRDVERFMIGIKTSLYRICMAVYPLMLTILTYLYEPLKTKD-YTQSGV
Qy      543  YVSGYLMFLGTFPPIGISLY-----KYRGTSETIKKAHFKSPSGMSPR
Db      570  YVFGMCL-----SAGVGVGLFWAIPAVRKOPSHGLMAR-IRKAFEPPLMNGPSDPQ
Qy      594  ERREMPKAE--AKALRQKMTSRVKNLWYSI 624
Db      622  TLKRYQLFVGEGNANALFRSS-----IWHKI 648

```

## RESULT 5

Q9V5B3 PRELIMINARY; PRT; 651 AA.

```

ID  Q9V5B3  PRELIMINARY; PRT; 651 AA.
AC  Q9V5B3  01-MAY-2000 (TREMBlrel. 13, Created)
DT  01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT  01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE  CG1698 protein.
GN  CG1698.
OS  Drosophila melanogaster (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC  Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX  Ephydroidea; Drosophilidae; Drosophila.
RX  NCBI_TaxID=7227;
RP  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=BERKLEY;
RX  MEDLINE=20196006; PubMed=10731132;
RA  Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA  Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA  George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA  Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA  Brandon R.C., Rogers Y.-H.C., Blake J.R., G. Chumpe M., Pfeiffer B.D.,
RA  Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA  Abril J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA  Bailew K.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA  Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA  Borokova D., Botchan M.R., Bouck J., Brodeur P., Brotler P.,
RA  Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA  Cherry J.M., Delcher A., Deng Z., Davenport L.B., Davies P.,
RA  de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA  Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA  Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA  Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA  Glisak A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA  Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA  Houslin D., Houston K.A., Howland T.J., Wei M.-H., Ideyama C.,
RA  Jaislin M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA  Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA  Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA  Mekulov G., Milshina N.V., Mobarry B., Murphy L., Muzny D.M., Moshrefi A.,
RA  Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA  Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA  Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA  Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA  Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA  Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA  Svirbas R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,
RA  Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA  Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA  Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA  Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

```

```

RA  Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT  "The genome sequence of Drosophila melanogaster."
RL  Science 287:2185-2195(2000).
DR  EMBL: AE003832; AAF58902.1;
DR  FlyBase: Fg00033443; CG1698.
DR  InterPro: IPR000175; Na/ntan_symport.
DR  Pfam: PF00209; SNF; 1.
DR  PRINTS: PR000176; NANUSMPORT.
DR  Prodom: PD000448; Na/ntan_symport; 1.
DR  PROSITE: PS0267; NA/NEUTROTAN_SYMPT_3; 1.
SQ  SEQUENCE 651 AA; 72070 MW; 30C81FA13C973285 CRC64;

```

Query Match 44.5%; Score 1492; DB 5; length 651;  
 Best Local Similarity 48.4%; Pred. No. 7.7e-111;  
 Matches 305; Conservative 96; Mismatches 191; Indels 38; Gaps 12;

```

Qy      11  ESSEKMEPKKSSQSLSPANNKALNDIDTLEAPP--ERWVMSNIEFLMSTIATNS 68
Db      38  ESSNGKKTLSDBGSGSAPPFGHDS-----PSGEAGQPKKRDNNIDIEFLMSTIATNS 91
Qy      69  VGLGNVWRFPPIAYONGGAFVYVYVLLNGKPPVYIECVYAGFSSRSNVKMSISPA 128
Db      92  VGLGNVWRFPPIALENGGAFVYVYVLLNGKPPVYIECVYIECVYIECVYIECVYIECV 151
Qy      129  MKGTGYAAGCCGTYLTYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 187
Db      152  MRGIGYQVLAATGVTYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 211
Qy      188  SVNNITNGTSSAQLYFRTVLOOSDGEGLAPVYVYVYVYVYVYVYVYVYVYVYVYVYV 247
Db      212  RATSLAGCGCTHSPHNITLRKASIDIDGIPSSVSLALALAVAYVYVYVYVYVYVYV 271
Qy      248  GKAAVYALPYPYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 307
Db      272  GKASYFLALPYPYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 331
Qy      308  TVCTGPPIIMSSYNGFRHNITRYDMVITLDFTFSLSCCTIFGLGLNAYELN-SEYGD 366
Db      332  ALFGNITIMAYNSRGHNITRYDMVITLDFTFSLSCCTIFGLGLNAYELN-SEYGD 391
Qy      367  VVAGGTSLAFISYPDAIAK-TFOPQLFSVLEFLMAYVGLGSSVALLSTNTLAMDAPF 425
Db      392  VVAGGTSLAFISYPDAIAK-TFOPQLFSVLEFLMAYVGLGSSVALLSTNTLAMDAPF 450
Qy      426  RVPYTVMSAMTSCGFLGLVYCTPGGOYILEVDHYGTFLVFCALSGLAGVMIYGL 485
Db      451  HKNMTVVGVIAVGFGLLTYTPGGOFLLLVDFGVFVALVALAIFELVLTAMIXG 510
Qy      486  ENULCDIDIEFMKKTGAWYRLCGVITPAIMTVTFVYALLASNNLVFGDNVYPTAGYS 545
Db      511  KRLCRDVERFMIGIKTSLYRICMAVYPLMLTILTYLYEPLKTKD-YTQSGVYVF 569
Qy      546  GYIMFLGTFPPIGISLY-----KYRGTSETIKKAHFKSPSGMSPRPRER 596
Db      570  GMCL-----SAGVGVGLFWAIPAVRKOPSHGLMAR-IRKAFEPPLMNGPSDPQTK 621
Qy      597  EMOPKAE--AKALRQKMTSRVKNLWYSI 624
Db      622  RYQLFVGEGNANALFRSS-----IWHKI 645

```

## RESULT 6

Q9VJRA PRELIMINARY; PRT; 639 AA.

```

ID  Q9VJRA  PRELIMINARY; PRT; 639 AA.
AC  Q9VJRA  01-MAY-2000 (TREMBlrel. 13, Created)
DT  01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT  01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE  BG-DS03431.1 protein (CG16161P).
GN  BG-DS03431.1 OR CG15279.
OS  Drosophila melanogaster (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC  Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

```



Query Match	43.9%;	Score 1472.5;	DB 5;	Length 639;
Best Local Similarity	47.8%;	Pred..No. 2.7e-109;		
Matches 302;	Conservative 97;	Mismatches 200;	Indels 33;	Gaps 11

[illegible][illegible]

RESULT 7			
ID	Q9NKA8	PRELIMINARY:	PRT; 739 AA.
AC	Q9NKA8;		
DT	01-OCT-2000 (TREMBLrel, 15, Created)		
DT	01-OCT-2000 (TREMBLrel, 15, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel, 21, Last annotation update)		
DE	Hypothetical 82.8 kDa protein.		
GN	BG:DS03431.1 OR CG15279.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_Taxid=7227;		
MI	[1]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN=BERKELEY;		
RC	MEDLINE=99403001; PubMed=10471707;		
RA	Asburner M., Mista S., Roote J., Lewis S.E., Blazer R., Davis T.,		
RA	Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,		
RA	Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,		
RA	Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., WhiteLaw K.,		
RA	Seinkner S., Rubin G.M.;		
RT	"An exploration of the sequence of a 2.9-Mb region of the genome of		
RT	Drosophila melanogaster: the Adh region.";		
RL	Genetics 153:1179-219(1999).		
RM	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY;		
RA	Seinkner S.E., Agbayani A., Arcaina T.T., Baxter E., Blazer R.G.,		
RA	Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,		
RA	Fafian D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,		
RA	Houston K.A., Hummasti S.R., Kaira K., Kearney L., Kim E., Lee B.,		
RA	Lewis S., Li P., Lomocan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,		

RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequelra A.,  
 RA Sechi H., Snir E., Svirska R.R., Wan K.H., Weinburg T., Zhang R.,  
 RA Zieran L.L., Rubin G.M.;  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AE003411, AAF44897.1;  
 DR FlyBase: FBgn028886; BG:DS03431.1.  
 DR InterPro: IPR000175; Na/ntiran\_symport.  
 DR Pfam: PF00209; SNF; 1.  
 DR PRINTS: PR00176; NANEUSMPORT.  
 DR ProDom: PD000448; Na/ntiran\_symport; 3.  
 DR PROSITE: PS50267; NA\_NEUTROTAN\_SYM\_3; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 739 AA; 82800 MW; 5C2FEFF22B5389FC5 CRC64;

Query Match 43.6%; Score 1461.5; DB 5; Length 739;  
 Best Local Similarity 42.2%; Pred. No. 2,4e-108;  
 Matches 309; Conservative 98; Mismatches 122; Indels 133; Gaps 13;

QY 21 RSSGSLPPANKKAL-DNIDDTLEAPPRMYSNNIEFLMSCIATSVGLAVWRPF 79  
 DB 19 RSTASTVEISTNSPALRDNSDDEAKYPEREATWGVGERLMSYGLVGNWREPF 78  
 QY 80 IAYONGGAFLPYVIVLLVKGPRVYLLGQFSSNSYKWSISPAK----- 130  
 DB 79 TALDNGGAFILPILVFLGKPYLLEMYTIGQFSRGYKVDLCPAMEYDSELON 138  
 QY 131 -----GTGYAAGCGYLYSYVYVGLCLYLLAMSPQATLPWA 169  
 DB 139 LHRSHVQLARSCPKYATTSIGVIGVISMTYVVALMGTLLVLESFSPLPWS 198  
 QY 170 ICGEPHEN-CYPS-----DPLLAAS----- 188  
 DB 199 ECREWESFCVASSLNGTSSLASPMDSLPQORPVASELVFLTYAALVATIGHYFI 258  
 QY 189 -----VNNTNGTSSAOL-----YF 203  
 DB 259 ESFENPLPMTSCRAEMKIHICNSAPDSNMSQLESNDORPNTMKSONDRVYTSSEMYF 318  
 QY 204 LRTVLOQSDGIEGLGAPIMVLYCLFETIAMLVGVYARGVSSGKAAYFLAPYVMI 263  
 DB 319 VKEVLHKEPNIEEGIGLPMNELVIGLFLWSCVFILIRGVKSSGKASYFLALPEYIMG 378  
 QY 264 TLETTITILPGANDGILFVTPQAKLLELGWYSANTQVFFSLTCTGTPIMSSYNGE 323  
 DB 379 VLVRAVATLPQSDIGIYFTRKPMQKILDPKVVAAVTCQFYSLSVCEGNIMYSSPNKF 438  
 QY 324 RHNIYRDAMIYVTLDTFTSELSGCTIGILGNLAYELNS-EVGVYVAGGTSIAFISYD 382  
 DB 439 GHNVHRAALVYTGDMTSLAGTTIGLILHAIHEIGTDIGSYV-KGGAGLAFISYD 497  
 QY 383 AIAKTFQ--POLSEVFLFMSVYLGSSVALLSTFNTLAMDAFPRVPTYMSAMTSCG 440  
 DB 498 AIAK-FKNLPOIFSVLEFELMFVYLGISNTIAMTSCSYTAIRDRPNFGQMCSSILLIAVS 556  
 QY 441 FLGLVYCTPGGQVITELVDHYGTEFLVLCATSELAGVWYIGLENLCLDIEMMLKKT 500  
 DB 557 FETGLMITTPGGQMLTVDFPGSMALVGLAEILTYITWITGTDRLCLDIEMMLGRVY 616  
 QY 501 GAVWRLCWYITPAIMTVEFYFALLASNNLVFGDNYVTPAGYSGYLMFLGMYFPIG 560  
 DB 617 GLVWRLCWYITPLIMVILLYFYATYQPLTY--NNIYYPKMSISIGMLITAFGLQPLW 675  
 QY 561 TGSRLKYRTGSETIKKAFHSKPSWGPSPRREREMOFAKAKALROKMTNSRYKHL 620  
 DB 676 MIYAVIVDPQGTGAKIRGAFTPKKNWGPSPDLLREQY-----HKEIEMELPKRGQGI 729  
 QY 621 WYSITGAYRRNI 632  
 DB 730 W-----AIAIKONI 737

RESULT 8  
 Q9V690

ID Q9V690 PRELIMINARY; PRT; 629 AA.  
 AC Q9V690:  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE CG8850 protein.  
 DE CG8850.  
 GN Drosophila melanogaster (fruit fly).  
 OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_Taxid=7227;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RP STRAIN=BERKELEY.  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amandalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pallos R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Mei M.H., Idegawa C.,  
 RA Jatta M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003822; AAF58358.1;  
 DR FlyBase: FBgn0033708; CG8850.  
 DR InterPro: IPR000175; Na/ntiran\_symport.  
 DR Pfam: PF00209; SNF; 1.  
 DR PRINTS: PR00176; NANEUSMPORT.  
 DR ProDom: PD000448; Na/ntiran\_symport; 2.  
 DR PROSITE: PS50267; NA\_NEUTROTAN\_SYM\_3; 1.  
 SQ SEQUENCE 629 AA; 69715 MW; F50871F7358BCD5B CRC64;

Query Match 42.2%; Score 1414; DB 5; Length 629;  
 Best Local Similarity 43.9%; Pred. No. 1.3e-104;  
 Matches 280; Conservative 113; Mismatches 209; Indels 36; Gaps 9;

QY 10 FESSEPKMEPKRSSQISLPPANKKALDNIDDTLEAPPE-----RWWS 55  
 DB 7 YOKLRNMGAQSGRCHPGSN-----DGISTVIYSAEGBELTINCAESESSESGORDWS 61  
 QY 56 NNIEFLMSCIATSVGLAVWRPFPIAVONGGAFLPYVIVLLVKGPRVYLLGKPYLLECYGQFS 115

Db 62 RGVEFLSCIALSVGLGNWRPFPALENGGAFLLPYVLLILGRPVYLLVLEIIQGS 121  
 QY 116 SRNSVKWYSPAMKNGVQAAGCGIILSYVYVIGLCIYLAMSQATLPAICOPEN 175  
 Db 122 SRCIRAFDMAPIMKRIAGQVYSTALATYVACIALMLTRIVASSEVLPMTYCLVEM 181  
 QY 176 -ENCVPSPDPTLAASVNNITNGTSSAQLYFLRTVLOOSDGI-EGGLGAPIVYLCLFIAM 233  
 Db 182 GKSCVATGANT-ANNDSTVGVSSALFETQVTLREPSIDNDGLSPSDVLCLLATW 240  
 QY 234 LAMFVYARCKVSSGKAATFLAPYVVMITTEFTITIIIPGATGILFEVTPQWAKLLEL 293  
 Db 241 VIIGTILSKIRSSGASYFLALFPYVIMVILLIRAVTLIPGANOGLVYFLKPGMSOLNP 300  
 QY 294 GWYSAVTOVFEFLVCTGPIIMFSSYNGRHNIVYDAMVITLDTFTSLSCCTFGL 353  
 Db 301 HWYVAITQMEFLALCFGLVWYASFPNDKNVHDVITITIDSLTSLACIITFGL 360  
 QY 354 GNATYELNSEGVGVAGTSLAFISYPAIAK-TPQPOLSVYFLFLLMSVGLIGSSVAL 412  
 Db 361 GNATFETNKRDISQVYKGGAGLAFISYPAIAKFKYLPQFAVLFPFLLVIGISNIGM 420  
 QY 413 LSTFNTLADAPFRVPTVYVMSAMTSCGFLIGVCTPGQVYLLVDHNGFVLYFCA 472  
 Db 421 ASAVVAVVADRFTLPHMLIYASASITIGLCGLVMTPGQVFLNLDVFGCTFIALVLA 480  
 QY 473 ISELAVEMLYGLENLCDIEFMLGKRTGAYWRLCGVITPAIMTVFFAYALLASNNLVE 532  
 Db 481 IAEVAVGVYGVKRICSDIEFMLNVKTSFYWMICAIYAPGLMFLVLYMLFSYEPY 540  
 QY 533 GDNVYVPIAGVYSGVYMLFLGVFVIGISLTKRTGTFSSETTIKAFHSKPSWGRSP 592  
 Db 541 -RGVOYPPVYVYAGVIMGLVQLDLPFMAVLYTIOQPKFTGSKFGLAQPANMPLQT 599  
 QY 593 RERREWMOPKAKALRQK---MNTSRVKLHWYSYG 626  
 Db 600 Q-----KFEAYIIHRRKREDFKSPRGCIYLFDNIFG 629  
 RESULT 9  
 Q9VSV2 PRELIMINARY; PRT: 593 AA.  
 AC Q9VSV2: 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE CG4476 protein.  
 GN CG4476.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydriidea; Drosophilidae; Drosophila.  
 NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Branton R.C., Rogers Y.H.C., Blazer R.G., Champs M., Pfeiffer B.D.,  
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Abmayant A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Bendandi J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes K., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ideyama C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kelchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mitsuhashi N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Rabinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RA "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003552; AAF50309.1; -  
 DR FlyBase: FBgn0035969; CG4476.  
 DR InterPro: IPR000175; Na/ntiran\_sympor.  
 DR InterPro: IPR000504; RNA\_rec\_mot.  
 DR Pfam: PF00209; SNF\_1.  
 DR PRINTS: PR00176; NANEOSMPORT.  
 DR ProDom: PD000448; Na/ntiran\_sympor; 1.  
 DR PROSITE: PS0267; NA\_NEUROTRAN\_SYMP\_3; 1.  
 DR PROSITE: PS00030; RRM\_RNP\_1; UNKNOWN1.  
 DR SEQUENCE 593 AA; 66043 MW; 26D7042ADB3DBA0 CRC64;  
 Query Match 41.4%; Score 1389.5; DB 5; Length 593;  
 Best Local Similarity 49.0%; Pred. No. 1.1e-102;  
 Matches 284; Conservative 89; Mismatches 188; Indels 19; Gaps 10;  
 QY 39 IDDTLEAPPRPFWNSNIEFLMGCINTSVGLGNWRPFPAAYONGGAFLLPYVYL 98  
 Db 3 LESSELSPPKARDKMGSSLEFLMGCIALSVGLGNWRPFPALENGGAFLLPYVYL 62  
 QY 99 LVGKPVYVLYCLEVAGOFFSSRNSVYKWSISPAKMGTYQAAGCGIILSYVYVIGLCIYLL 158  
 Db 63 VVGKPIYVWEMLLGQFSSRGIVQVDFAPIMRGVQAQLALGVATYAYASMALTLRYF 122  
 QY 159 AMSFQATLPWALCOPEM-ENCVPSPDPTLAASVNNITNGTSSAQLYFLRTVLOOSDGI 217  
 Db 123 FDSFASSELPMWSCREMDGCVASAGGQPLQGLSRNFSSTOTLQRIIVNETDSEEG 182  
 QY 218 LGAPITVLYLCLFIAMLVFVGVARGVSSGKAAYFLALFPYVVMITTEFTITIIIPGAND 277  
 Db 183 IGYPGSLALMGLIMLVITLIIINGVSSGKAAYVLAFLPYVWIFILVRAITLPGAYD 242  
 QY 278 GILFVITPQWAKLLELGVYSAVTOVFEFLVCTGPIIMFSSYNGRHNIVYDAMVITTL 337  
 Db 243 GVMYFLTPQWEXLELPQVYVNAVTVQVFEFLVACFVLIIMYSSYNGRHNIVYDAMVITTL 302  
 QY 338 DFTFELSGLCTFGLIGLNLAYELNS-ENGVDVAGAGTSLAFSYDATALK-TPQQLFSV 395  
 Db 303 DFTFELSGLVITFGLIGLNLAYELNS-ENGVDVAGAGTSLAFSYDATALK-TPQQLFSV 361  
 QY 396 LEFLMSVGLIGSSVALLSTENTLAMDAPFRVPTVYVMSAMTSCGFLIGVCTPGQVY 455  
 Db 362 LEFLMFLMGVSGNVGSMCIIMTVLKDQFVNVKMIIVVLSVIGFLVGLIYTPGGCHI 421  
 QY 456 LEIVHNGCTVLYFCALSELVAGVYVIGLENLCIDIEFMLGKRTGAYWRLCGVITPAI 515  
 Db 422 ITLMDHEVTVSVLSALFELIAGVITGRLODADAYMINKTSNYRNCWSIYVPLV 481  
 QY 516 MTVEFYVALLASNNLVE-GDNV--VYPYAG-VSGYMLFLGTMFVPIGIFSLYKRYTG 571  
 Db 482 MLVILVYSLMLMRPLSTVNGQFPLVYRVYGVCSGCT---IGOLF--YMAQIANFKOPKG 536

OY 572 TFSITKKAFHSPKSPRPREREMOKFAKALROK 611  
 DB 537 SLKSRINNSIKPHSDMCPSPKLMYOMF-----LRNK 570

## RESULT 10

0877E1 PRELIMINARY; PRT; 615 AA.

AC 0877E1; (Tremblrel. 21, Created)  
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Amino acid transporter.  
 OS Aedes aegypti (Yellowfever mosquito).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;  
 OC Aedes.  
 OX NCBI\_TaxID=7159;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Dasher M.K., Kohn A.B., Harvey W.R., Stevens B.R.;  
 RT "AENART, A Novel Amino Acid Transporter from Larval Aedes aegypti  
 RT Midgut.";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF369383; AAA13400.1;  
 SQ SEQUENCE 615 AA; 68985 MW; E14A5ED2E809A139 CRC64;

Query Match 36.98; Score 1238.5; DB 5; Length 615;  
 Best Local Similarity 46.08; Pred. No. 1.3e-90;  
 Matches 267; Conservative 79; Mismatches 173; Indels 61; Gaps 12;

OY 76 RPEPIAYONGGATFVYVILLVGRPVYLLVGLGQSSRSNVKWSISPAKKGTYA 135  
 DB 44 RPEPIAYONGGATFVYVILLVGRPVYLLVGLGQSSRSNVKWSISPAKKGTYA 103  
 OY 136 QACGCGYILSTYVYVIGLCLYLLMSPOATLPAICPEW-ENCYSPDTLASV---NN 191  
 DB 104 QLVQTVSVYVYVLLVGLLTHYIPASFASELPMATCKDNMADNCV--DSSLVNMEDSN 161  
 OY 192 ITNG--TSSAQLYELRTVYVLLVGLGQSSRSNVKWSISPAKKGTYA 249  
 DB 162 STSQKQYSSQIYVLDLYLAKKSDIDGICAPPMKLLMLLMAVYVFLVLYKGVKSSG 221  
 OY 250 AAYFLALFPYVYVITLFTTILPGATDGLFVYVTPQMAKLELGVYSAVTVQFSLTV 309  
 DB 222 VAYFLALFPYVYVITLFTTILPGATDGLFVYVTPQMAKLELGVYSAVTVQFSLTV 281  
 OY 310 CTGPIIMSSYNGRHNHYRDAMVITLDTFTSLGCTTIGTIGLNAVELNSEGVYV 369  
 DB 282 GMSGILMFSSYNNHNNHYRDAMVITLDTFTSLGCTTIGTIGLNAVELNSEGVYV 341  
 OY 370 AGGSLAFLSYPDIAK-TEPOLFSVFLFMSVGLIGSSVALLSTFNTLAMDAPRPV 428  
 DB 342 KSGGLAFLSYPDIAKFDIYVQFVFLFMSVGLIGSSVALLSTFNTLAMDAPRPV 401  
 OY 429 TVYSAMTSCGGLGLVY--CTPGGQYILELVHGYSTFVLCASISLAGVWYIGLE 486  
 DB 402 YMHLLALISTIGFTGLVYHILVANGYSIMSTTE--AFSDLLALIEWVAIWIYGLD 458  
 OY 487 NLCDIEFMKKGAVWRLCMGVITPAIMTTFVYVALLANNLVFGD----- 534  
 DB 459 NWCNDIEFMVORVGLY-----VAANLGSNNATFHCORHFLVLEK 501  
 OY 535 ----NYVPTAGVYSGVLMFLGATFVYVIGLGLYKRTGTFSE-----TIKAFHSP 585  
 DB 502 PTYGGQYVREGALIGIFVIFGLIYVLMVWIV---TGDSTRKESWGSISKAAPSS 557  
 OY 586 SWGPRSPREREMOKFAKALRO-----KMTSRVXKL 620  
 DB 558 QMPCTEHTTKRAMLKYKEAKSRDDLIYAKNHSSIVAKL 597

RESULT 11

091202  
 ID 091202; PRELIMINARY; PRT; 791 AA.

AC 091202; (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Glycine transporter type 2.  
 GN GLYT2.

OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA STRAIN-BALB/C; TISSUE-BRAIN;  
 RT Liu Q.-R., Li Q.-F.;  
 RL "Cloning and expression of mouse glycine transporter 2 (Glyt2)."  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF411042; AAL17054.1;  
 DR InterPro: IPR001920; Asp/Glu\_race.  
 DR Pfam: PF00209; Snp. 1.  
 DR Prodom: PD000448; Na/utran\_symport. 2.  
 DR PROSITE: PS00923; ASP\_GLU\_RACEMASE\_1; UNKNOWN\_1.  
 DR PROSITE: PS00610; NA\_NEUROTRAN\_SYM\_1; UNKNOWN\_1.  
 DR PROSITE: PS00754; NA\_NEUROTRAN\_SYM\_2; 1.  
 DR PROSITE: PS50267; NA\_NEUROTRAN\_SYM\_3; 1.  
 SQ SEQUENCE 791 AA; 87029 MW; 8AABB600D9EAEF CRC64;

Query Match 32.58; Score 1089.5; DB 11; Length 791;  
 Best Local Similarity 37.68; Pred. No. 1.4e-78;  
 Matches 225; Conservative 105; Mismatches 229; Indels 39; Gaps 9;

OY 35 ALNDIDTDLEAPPRERWVNNIEFLMCIATSVGIGWVRRPEPIAYONGGATFVYV 94  
 DB 169 ATTTTPEDEGDEKARKGNSSKLDPLTSVAGVAGLGNWRRPPIYAFONGGATFVYV 228  
 OY 95 IVLLVGRPVYLLVGLGQSSRSNVKWSISPAKKGTYAQAAGCGYILSTYVYVIGLC 154  
 DB 229 MMLALGLPIFLFVLSIGQFASOGPVSWKAIPALOGCGIAMLISVLAIVYVITCYT 288  
 OY 155 LYTLAMSFQATLPAICPEW-----ENCYSPD-----PRT 185  
 DB 289 LFTLFAFVSVLPMGSCNNMWNPECKDKTKLLDSCVIDDHKIQIKNSPCMTAVPNL 348  
 OY 186 A-ASVNNITNGT--SSAQLYELRTVYVLLVGLGQSSRSNVKWSISPAKKGTYA 242  
 DB 349 TVNFTSQTKTFVSGSEEFKFKVLAISGIEP-GEIRPLAFCLFLAMVIVYASLAK 407  
 OY 243 GYSSGKAATFLALFPYVYVITLFTTILPGATDGLFVYVTPQMAKLELGVYSAVTV 302  
 DB 408 GISSGKAVVYVFTFEPYVYVILLIRGVTLPGAGAGIWTYITRKWEKLTDTATVKDAATQ 467  
 OY 303 VEFSLVCTPIIMFSSYNGRHNHYRDAMVITLDTFTSLGCTTIGTIGLNAVELNS 362  
 DB 468 IFFSLSAMGGLITLSSYNNHNNHYRDAMVITLDTFTSLGCTTIGTIGLNAVELNS 527  
 OY 363 EGVGVAGGTSIAFLSYPDIAKTFPOLFSVFLFMSVGLIGSSVALLSTFNTLAMD 422  
 DB 528 NIENVADO-SPGIAFVYVPEALTRLPSPMAIFFLMLTLTLDITFATIEYIVTISID 586  
 OY 423 AFRPVYV 479  
 DB 587 EFRPVYV 646  
 OY 480 FVYV 539  
 DB 647 SYVY 705  
 OY 540 TAGVY 597  
 DB 706 NMSVY 762

Query Match	31.8%	Score 1067.5	DB 5	Length 415
Best Local Similarity	52.9%	Pred. No. 3.7e-77		
Matches 199	Conservative 73	Mismatches 97	Indels 7	Gaps 5

Query Match	30.7%	Score 1030	DB 11	Length 638
Best Local Similarity	38.5%	Pred. No. 6e-74		
Matches 230	Conservative 93	Mismatches 225	Indels 50	Gaps 14
OX	NCB1_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-DDY;			
RA	Ugwa S.;			
RT	"Mouse Colonic System B0+ Amino Acid Transporter."			
RL	Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB033285; BAA94300.1; -			
DR	MGD; MGI:1890216; S1C6A14.			
DR	InterPro: IPR000175; Na/ntan_sympot.			
DR	Pfam; PF00209; SNE; 1.			
DR	PRINTS; PR00176; NAMEUSMPOT.			
DR	ProDom; PD000448; Na/ntan_sympot; 2.			
DR	PROSITE; PS00610; NA_NEUTROTRAN_SYMP.1; 1.			
DR	PROSITE; PS50267; NA_NEUTROTRAN_SYMP.3; 1.			
SO	SEQUENCE 638 AA; 71455 MW; 505A7B2DF7E7612 CRC64;			
QY	47	EPERAMWNSNIEFLMSCIATSVGLGNWRRPPIAYQNGCAFIPYIVILLNKPPYY	106	
Db	32	ENORGNMKSDDLVSNGYAVGANGWRRPPIYLTGGAFLPIALMALGLPFL	91	
QY	107	LECVAGOFSSSNYSKWSISIPAMKGTGAQAAGCGYILSYVVICGLCLYLIAMSFOATL	166	
Db	92	LECSLGGFASLGPVSVRIPLPLFGVGITIMVLISVFVAIYTNVIATSLYILFASFQSVL	151	
QY	167	PWALICQPEW--ENCVPSPDPIILASV-----NNIT--NGTS-----SA	199	
Db	152	PWANC--SWADENCSPRPVYGVCSVIGAGEMFNISVWNTNNLCLNGSEVFRPQLPS	210	
QY	200	QLYVLRVLYOOSDIEGLGAPIMVLYLCLFIAMVMGVAVRGYSGSKAAYFLALPPY	259	
Db	211	EYWDKTYLORSSGMD--ETGVVYVYLALCLLMLVIGALFLKGIKSSGKVVYFTALPPY	269	
QY	260	VVMITLFTITTLIPGATDGLIEFVTPQ--NAKLELGWYSAVTVQVPSLTVCCTPIIME	317	
Db	270	VLLILLIRGATLEGASKGISYIGAGSNFTRKREABEVWKAQAQIFSLSVAMGVAL	329	
QY	318	SSVNGFPHNITRYDAWITTTDTFSPSLSGCTIFGLGLMAELANSEGVDDVAGAGTSIAF	377	
Db	330	SSVKNFNNNCSDAIIVCLNGLISVPAAGFAIFSLGMAHISGKEVSVV--KSGFDLAF	388	
QY	378	ISYDAIAKTFEPOPFLSYFLFLMSVYIGGSAVLLSTFENFLAMDAP-----RVPTVY	431	
Db	389	IAYEALAAQLPAGPFWGSLFFFLMLTLGLDSQPSASIEITTTTFODLPFKAKRRARVP--	445	
QY	432	MSAMTSCGCGFLGLVYCTPGGQYILEVDHYGTFELVFCALISELAGVWITGLNCLD	491	
Db	446	ITLGCCLLIEGLICLYTQAGIYVWHLDHFCACMGWILLIAILEIAGIIMYIGGNRIED	505	
QY	492	IEPMLGKGTGAY--WFLCGVITPRAMTYFEFFALLASNNLVYCGDNVYFTAGVSSYL	548	
Db	506	IEEMITGAKRMIFWLMWRACWFLVTPILLSALVLSVLFKFNRPDAD--IPYDWGVALGWC	564	
QY	549	MLELGMFVFDIGLGFSLYK--YRTGFSEITIKAFHSPKSPGSPRREREMQFKAE	604	
Db	565	MIIFCIIMWPI---MAIKIIVQAGSNILQRIISCRPASNNGPFLERKRGERYDMAE	619	
RESULT 14				
Q8WPM9	PRELIMINARY;	PRT;	674 AA.	
AC	OBWPM9;			
DT	01-MAR-2002 (Tremblrel. 20, Created)			
DT	01-MAR-2002 (Tremblrel. 20, Last sequence update)			
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)			
DE	Similar to glycine transporter.			
GN	BAC001.6.			
OS	Oikopleura dioica.			
OC	Eukaryota: Metazoa: Chordata: Urochordata: Appendicularia.			



Db	506	ITEMIGAKRMIFWLMWRACWFIPTPILLSAILWMSLYKFRHPDYAD-IPYDNGVALGWC	564
Qy	549	MLFLGMTFVPICIGIFSLYK--YRTGFSEETIKKAFHSKPSWGRSPRERREMMQPKAE	604
Db	565	MIIFCIWIPI---MAIKIYQAEQNIILORITSCCIPASNMGPYLEKHERGERYRDMAE	619

Search completed: July 21, 2003, 09:27:47  
Job time : 91 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: July 21, 2003, 09:24:45 ; Search time 18 Seconds

(without alignments)  
1034.706 Million cell updates/sec

Title: US-09-991-458-2

Perfect score: 3354  
Sequence: 1 MNOGVNNGFESSSEPKMEPK.....TSRYKHLMTGAYRRNIN 633Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/1aa/PCUTUS.COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1084.5	32.3	797	2	US-08-700-013B-21	Sequence 21, Appl
2	1082.5	32.3	797	4	US-09-182-728A-2	Sequence 2, Appl
3	1082.5	32.3	797	4	US-09-795-232-2	Sequence 2, Appl
4	1080.5	32.2	797	2	US-08-700-013B-19	Sequence 19, Appl
5	1080.5	32.2	797	4	US-09-191-468-122	Sequence 122, App
6	1080.5	32.2	797	4	US-09-191-468-124	Sequence 124, App
7	1076.5	32.1	797	4	US-09-191-468-120	Sequence 120, App
8	1076.5	32.1	799	2	US-08-700-013B-27	Sequence 27, Appl
9	1030.5	30.7	667	1	US-07-879-617A-8	Sequence 8, Appl
10	1030.5	30.7	667	1	US-08-753-985-8	Sequence 8, Appl
11	1017	30.3	635	1	US-07-879-617A-9	Sequence 9, Appl
12	1017	30.3	635	1	US-08-753-985-9	Sequence 9, Appl
13	1011	30.1	599	1	US-08-301-722A-5	Sequence 5, Appl
14	1008	30.1	599	1	US-08-295-814E-11	Sequence 11, Appl
15	1008	30.1	599	1	US-08-240-783B-4	Sequence 4, Appl
16	1008	30.1	599	3	US-09-084-813-4	Sequence 4, Appl
17	1008	30.1	599	4	US-09-343-361-11	Sequence 11, Appl
18	1008	30.1	599	5	PCT-US92-09662-4	Sequence 4, Appl
19	991	29.5	630	1	US-07-959-943-11	Sequence 11, Appl
20	983	29.3	599	1	US-07-879-617A-10	Sequence 10, Appl
21	983	29.3	599	1	US-08-753-985-10	Sequence 10, Appl
22	981	29.2	630	1	US-07-959-943-9	Sequence 9, Appl
23	978.5	29.2	632	4	US-08-295-814E-10	Sequence 10, Appl
24	978.5	29.2	632	4	US-09-343-361-10	Sequence 10, Appl
25	978.5	29.2	632	5	PCT-US93-01959-10	Sequence 10, Appl
26	973	29.0	627	1	US-08-295-814E-4	Sequence 4, Appl
27	973	29.0	627	4	US-09-343-361-4	Sequence 4, Appl

28	973	29.0	627	5	PCT-US93-01959-4	Sequence 4, Appl
29	971.5	29.0	607	1	US-07-959-943-7	Sequence 7, Appl
30	968.5	28.9	653	1	US-07-782-298-2	Sequence 2, Appl
31	967	28.8	687	3	US-08-834-467-2	Sequence 2, Appl
32	967	28.8	687	4	US-09-336-177-2	Sequence 2, Appl
33	965	28.8	602	1	US-08-295-814E-2	Sequence 2, Appl
34	965	28.8	602	4	US-09-343-361-2	Sequence 2, Appl
35	965	28.8	602	5	PCT-US93-01959-2	Sequence 2, Appl
36	962	28.7	627	1	US-08-291-299-10	Sequence 10, Appl
37	962	28.7	627	5	PCT-US95-10579-10	Sequence 10, Appl
38	960.5	28.6	607	1	US-07-879-617A-12	Sequence 12, Appl
39	960.5	28.6	607	1	US-08-753-985-12	Sequence 12, Appl
40	958	28.6	614	1	US-08-291-299-7	Sequence 7, Appl
41	958	28.6	614	5	PCT-US95-10579-7	Sequence 7, Appl
42	953.5	28.4	621	1	US-08-295-814E-6	Sequence 6, Appl
43	953.5	28.4	621	4	US-09-343-361-6	Sequence 6, Appl
44	953.5	28.4	621	5	PCT-US93-01959-6	Sequence 6, Appl
45	949.5	28.3	620	1	US-08-301-722A-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-08-700-013B-21  
Sequence 21, Application US/08700013B  
Patent No. 5919653

## GENERAL INFORMATION:

APPLICANT: Albert, Vivian R.  
APPLICANT: Kowalski, Leslie R.Z.  
APPLICANT: Borden, Laurence A.  
APPLICANT: McKelvy, Jeffrey F.  
TITLE OF INVENTION: Human Glycine Transporter  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:

ADDRESS: Dechert Price & Rhoads  
STREET: 997 Lenox Drive, Building 3, Suite 210  
CITY: Lawrenceville  
STATE: NJ  
COUNTRY: USA

ZIP: 08543  
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/700.013B  
FILING DATE:

CLASSIFICATION: 536  
PRIOR APPLICATION DATA:

APPLICATION NUMBER:  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Bloom, Allen

REGISTRATION NUMBER: 29,135  
REFERENCE/DOCKET NUMBER: 317743-108

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 609-520-3214

TELEFAX: 609-520-3259  
TELEX:

INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:

LENGTH: 797 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear

US-08-700-013B-21

Query Match 32.3% Score 1084.5; DB 2; Length 797;  
Best Local Similarity 38.2% Pred. No. 8.5e-94;  
Matches 221; Conservative 106; Mismatches 213; Indels 39; Gaps 10;

```
QY 54 WSNNIEFLMSCIATSVGLGNVWREFFPIAYONGGAFVYVYVLLVGRPVYVECVLQ 113
D 194 WSKLDFILSMVGYAVGLGNVWREFFPIAYONGGAFVYVYVLLVGRPVYVECVLQ 253
QY 114 FSSNSVYKVMISIPAMKGTGAOAGCGYILSYVVICGLCLYTLAMSFQATLPMAICOP 173
D 254 FASOGPVSWKAIPLDGGCGIAMLINSVLAIVYVVICYLLFYLFAFSFVSLPWGSCNN 313
QY 174 EM-----ENCVPSD--PTL-----AASVNNIT--NGTSSA-----OL 201
D 314 PMNTPECKDKTKLLDSCVISHDKIQKNTFCMTAPVNTVMNFTSQANKTVSGSEE 373
QY 202 YFLRTVYQOOSDGIAGIAPYVYVLCFLAMLVFQVAVGKSSGKAAYFLALFPYV 261
D 374 YFKYFVLKISAGIYYP-GEIRMPALCLFLAMVIVYASLAKGITSKVVYFATFPYV 432
QY 262 MITFIFTIILPGATDGLFEPVPMQKLELGWYSAYVOVPSLVCCGPIIMESSYN 321
D 433 LVILLINGVILLPGAGAIWIFITPKMEKLDATYWKDAATQIFSLAAMGGLTTLSSYN 492
QY 322 GFRHNIYRDAMVITTLDTFTSFLSGCTIFGILGNLAVELNSEVDVYAGGTSIAFTSY 381
D 493 KFHNNCYRDLIVCTNSATSIFFAGFVIFSVIGFMANERKVNIENTVADQ--GPGIAFYVP 551
QY 382 DAIKTOPOLFVFLFLMSVLCIGSSVALLSTFNFLAMDAFPRTVYMSANT---CS 438
D 552 EALTRPLSPFWALIFELMLLTGLDTMFATIEIVTISIDPEPKYLRTHKPVFTLCCV 611
QY 439 CGFLGLVYCTPGGOYILEVDHYSGTEFLVFCALISLAVGFWYGLNCLDIEFMTGK 498
D 612 CFPIFMGPMTTGGIYMFOLVDYTAASYALYIATLFEVLGISTYVIGQRCEDEMTIGF 671
QY 499 KTGAVWRLCWGITPAIMTTFEYVALLASNNLVFGDNVYVYVYVYVGYVLMFLGTMFVP 558
D 672 QPNIFWKCWAFFYPTLITFELCFEFYQWEPMTYG--SYRYPNMSMVLGWLMLACSVIWP 730
QY 559 IGIGFSLXKXRTGTFESTIKKAFHSKPSNGPRSPRERE 597
D 731 IMFVYIKMH-LAPGRFIERLKLACSPQPDWGPFLAQHNGE 768

RESULT 2
US-09-182-728A-2
; Sequence 2, Application US/09182728A
; Patent No. 6238883
; GENERAL INFORMATION:
; APPLICANT: BROWN, ANTHONY
; APPLICANT: CHAPMAN, CONRAD GERALD
; APPLICANT: GLOGER, ISRAEL SIMON
; APPLICANT: EVANS, JOANNE RACHEL
; APPLICANT: CAIRNS, WILLIAM
; APPLICANT: HERDON, HUGH
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30176
; CURRENT APPLICATION NUMBER: US/09/182,728A
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 9818890
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Ver 3.0
; SEQ ID NO 2
; LENGTH: 797
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-182-728A-2

Query Match 32.3%; Score 1082.5; DB 4; Length 797;
Best Local Similarity 38.3%; Pred. No. 1.3e-93;
Matches 222; Conservative 105; Mismatches 213; Indels 39; Gaps 10;
```

```
QY 114 FSSNSVYKVMISIPAMKGTGAOAGCGYILSYVVICGLCLYTLAMSFQATLPMAICOP 173
D 254 FASOGPVSWKAIPLDGGCGIAMLINSVLAIVYVVICYLLFYLFAFSFVSLPWGSCNN 313
QY 174 EM-----ENCVPSD--PTL-----AASVNNIT--NGTSSA-----OL 201
D 314 PMNTPECKDKTKLLDSCVISHDKIQKNTFCMTAPVNTVMNFTSQANKTVSGSEE 373
QY 202 YFLRTVYQOOSDGIAGIAPYVYVLCFLAMLVFQVAVGKSSGKAAYFLALFPYV 261
D 374 YFKYFVLKISAGIYYP-GEIRMPALCLFLAMVIVYASLAKGITSKVVYFATFPYV 432
QY 262 MITFIFTIILPGATDGLFEPVPMQKLELGWYSAYVOVPSLVCCGPIIMESSYN 321
D 433 LVILLINGVILLPGAGAIWIFITPKMEKLDATYWKDAATQIFSLAAMGGLTTLSSYN 492
QY 322 GFRHNIYRDAMVITTLDTFTSFLSGCTIFGILGNLAVELNSEVDVYAGGTSIAFTSY 381
D 493 KFHNNCYRDLIVCTNSATSIFFAGFVIFSVIGFMANERKVNIENTVADQ--GPGIAFYVP 551
QY 382 DAIKTOPOLFVFLFLMSVLCIGSSVALLSTFNFLAMDAFPRTVYMSANT---CS 438
D 552 EALTRPLSPFWALIFELMLLTGLDTMFATIEIVTISIDPEPKYLRTHKPVFTLCCV 611
QY 439 CGFLGLVYCTPGGOYILEVDHYSGTEFLVFCALISLAVGFWYGLNCLDIEFMTGK 498
D 612 CFPIFMGPMTTGGIYMFOLVDYTAASYALYIATLFEVLGISTYVIGQRCEDEMTIGF 671
QY 499 KTGAVWRLCWGITPAIMTTFEYVALLASNNLVFGDNVYVYVYVYVYVGYVLMFLGTMFVP 558
D 672 QPNIFWKCWAFFYPTLITFELCFEFYQWEPMTYG--SYRYPNMSMVLGWLMLACSVIWP 730
QY 559 IGIGFSLXKXRTGTFESTIKKAFHSKPSNGPRSPRERE 597
D 731 IMFVYIKMH-LAPGRFIERLKLACSPQPDWGPFLAQHNGE 768

RESULT 3
US-09-795-232-2
; Sequence 2, Application US/09795232
; Patent No. 6426405
; GENERAL INFORMATION:
; APPLICANT: Anthony M. Brown
; APPLICANT: Conrad Gerald Chapman
; APPLICANT: Israel Simon Gloger
; APPLICANT: Joanne Rachel Evans
; APPLICANT: William Cairns
; APPLICANT: Hugh Jonathan Herdon
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30176-D1
; CURRENT APPLICATION NUMBER: US/09/795,232
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 9818890.7
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 797
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-795-232-2

Query Match 32.3%; Score 1082.5; DB 4; Length 797;
Best Local Similarity 38.3%; Pred. No. 1.3e-93;
Matches 222; Conservative 105; Mismatches 213; Indels 39; Gaps 10;
```

```
QY 54 WSNNIEFLMSCIATSVGLGNVWREFFPIAYONGGAFVYVYVLLVGRPVYVECVLQ 113
D 194 WSKLDFILSMVGYAVGLGNVWREFFPIAYONGGAFVYVYVLLVGRPVYVECVLQ 253
```

```
QY 54 WSNNIEFLMSCIATSVGLGNVWREFFPIAYONGGAFVYVYVLLVGRPVYVECVLQ 113
D 194 WSKLDFILSMVGYAVGLGNVWREFFPIAYONGGAFVYVYVLLVGRPVYVECVLQ 253
```

[illegible]

```

/ TELEFAX: 609-520-3259
/ TELFX:
/ INFORMATION FOR SEQ ID NO: 19:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 797 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ US-08-700-013B-19

Query Match      32.2%; Score 1080.5; DB 2; Length 797;
Best Local Similarity 38.2%; Pred. No. 2e-93;
Matches 221; Conservative 106; Mismatches 213; Indels 39; Gaps 10;

OY 54 MSNIEFLMSCDIAISVGLGNWRRPEFIYONGGAFILPYYIVLLVCKPYYIECVLGO 113
    :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 194 WSKRIDFLISWCAVAGVGNWRRPFYLFQNGGAFILPYLMMLALAGLPFFLEVLQ 253
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY 114 FSSRSYKWSISIPAMKGTGYAQAAGCGYILSYVVYIGCLLYLAMSFOATLPMAICOP 173
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 254 PASQGVSWKRAIPALQCGIAMLIIISVLAIIYVNIICYTFLEAFSVSLPWGSCNN 313
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY 174 EW-----ENCPVD-PTL-----AASVNNIT--NGTSSA-----QL 201
    -:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 314 PWNTPCKDKTKRLLDSCVISDHPKIQIKNSTFECTATPDPNTWMPFTSQANKTFVSGSE 373
    -:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY 202 YFLRTVLQOSDIEGGLGAPIMYLIVLCLEFIAMLVGVVARGVSKSRAAYLALFPYV 261
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 374 YFKYFVLKISAGIEP-GEIRNPLALCLFLAMVIVYASLAKIKTSKVVYFTATFPYV 432
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY 262 MITEFITITIIIPGATDGLFVYTPQMAKLLGVWYSAVTOVFESLIYCTGPIMFSSYN 321
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 433 LVILIRSVTLPGAGAGIWFYFETPEWKEKLTATVWKDAATQFFSLSAAMGGLITLSYN 492
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY 322 GFRNHYDAMVITLDTFTSPISCTIFGLIGNLAYELNSVGVGAGTSLAIFSP 381
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 493 KFHNNKYDITLVICTNSATSIFFAGFVIFSVIGPMANERKVIENVADQ-GGCIAFYVP 551
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY 382 DAIAKTOPQLFSLVFLPLMSVYLIGSSVALLSTFNTLAMDAPRVPIVYNSAMT---CS 438
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 552 EALTFLPLSPFWAIIFFMLILLTGLDTMFATIEETIVTSSIDSEFPKYLRTKHKVPFTLGC 611
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY 439 CGEILGLVYCPGGOYIILEVDHVGSTLYLFCALISELAGVWYIGLEMLCIDIFPMIGK 498
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 612 CFFLMGPFMTIYGGIYFMFOVDYTAASALVITAIIFELGVISYVGLQFPCDIEEMIGF 671
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY 499 KTGAWRLCQWGTIVTAIMTVEFYVALLASNNLVFSDNYYPRAGVSGYIMFLGTETVP 558
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 672 QPNIFWAKCMAFVPTLITLFLICCFEYQWEPRTYG-SYRYPRMSVVLGMLMLACSVIWP 730
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY 559 IGIGFSLYKRTGTSETIKKAFHFKPSKGRPSPPRE 597
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 731 IMFVIMH-LAPGRFIERLKIVCSPPQDWGPELACHRGE 768
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 5
US-09-191-468-122
/ Sequence 122, Application US/09191468A
/ Patent No. 6416975
/ GENERAL INFORMATION:
/ APPLICANT: Callagher, Michael J.
/ APPLICANT: Burgess, Loyd R.
/ APPLICANT: Brunden, Kurt R.
/ TITLE OF INVENTION: Human Glycine Transporter Type 2
/ FILE REFERENCE: 123110501
/ CURRENT APPLICATION NUMBER: US/09/191,468A
/ NUMBER OF SEQ ID NOS: 124
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO: 122
/ LENGTH: 797
/ TYPE: PRT
/ ORGANISM: Human

```



DB 307 PMGSCNNPMTPECKDKTKLLDSCVISHPKIQIKNSTFCMTAVPNTVNTSLANKT 366  
QY 197 --SSQLYELRTVLOQSDIGEGGAPLWVLCFLTAMLVGVAVGVSSGKAATFL 254  
DB 367 FVSGSEEFKYFVKISAGIEP-GEIRMPALCLFLAMVIVASLAGITSGKVYFT 425  
QY 255 ALFPVAVITFTITIIIPGATDGLFEVTPQMAKLELGYVSAVQVPSLAVCTPI 314  
DB 426 ATFPYAVVILLIGVILPGAGAGIWTTPKMEKLDATVWKDAATQIFPSLSAAMGL 485  
QY 315 IMSSSYNGRHRNIYRDAMITVTLDTFTSLSGCTIFGILNLAELNSEVDVAGAGTS 374  
DB 486 IITLSYKFNHNCRDILVICTNSATSIFFAGVYFVSIGFMANERKNINENADQ-CPG 544  
QY 375 LAFISYPAIAKTQPOLSVLFFLMSVIGSGSVALLSTFNLTAMAPRVPVYMSA 434  
DB 545 IAFVYFPAALRLPLSPFWAITFFLMLTLTGIDTFASIEFTIVSISEPFLYRTHPV 604  
QY 435 MT--CSCGFLGLVYCTPGGOYILELVHDYGFELVFCALISELAGVFWIYGLNLCD 491  
DB 605 FTLGCCICFFIMGPFMTQGGIYMFQVDTYASALVIAIFELVGSIVYGLQRCED 664  
QY 492 IEFMLGKTGAVWRLCGVITTPAINTVFFYALLASNNLVGDNYVTPAGVSGYLMF 551  
DB 665 IEMMIGFOPNIFWKVCMAFVPTLITFLICFSFYQWEMETYG-SYRPNMSKVLGMLMA 723  
QY 552 LGMFPVPIGIFSLKYRTGFSEETIKKAFKSPKSPGRSPRERE 597  
DB 724 CSVIMPIFMVYIKMH-LAPGRFIERKLKVCSPQDMPGLAQHGE 768

## RESULT 8

US-08-700-013B-27  
Sequence 27, Application US/08700013B  
Patent No. 5919653

GENERAL INFORMATION:  
APPLICANT: Albert, Vivian R.  
APPLICANT: Kowalski, Leslie R.Z.  
APPLICANT: Borden, Laurence A.  
APPLICANT: McKelvy, Jeffrey F.  
TITLE OF INVENTION: Human Glycine Transporter  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dechert Price & Rhoads  
STREET: 997 Lenox Drive, Building 3, Suite 210  
CITY: Lawrenceville  
STATE: NJ  
COUNTRY: USA  
ZIP: 08543  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/700,013B  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Bloom, Allen  
REGISTRATION NUMBER: 29,135  
REFERENCE/DOCKET NUMBER: 317743-108  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 609-520-3214  
TELEFAX: 609-520-3259  
TELEX:  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 799 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-700-013B-27

Query Match 32.1%, Score 1076.5; DB 2; Length 799;  
Best Local Similarity 37.5%; Pred. No. 4.8e-93;  
Matches 224; Conservative 107; Mismatches 228; Indels 39; Gaps 10;

QY 35 ALDNIIDDDLEAEPERPERVNNIEFLMSCATISGVGNWRPFPIYONGGAFVLPV 94  
DB 177 ATTIPEDGCDENKARGMSSKIDFLISMGVAVGNWRPFPLAFONGGAFLLPYL 236  
QY 95 IYLLVGRPVYLLCEVIGQFSSRNSVYKWSISPAKMGTYQAAGCYLLSYVYIGLC 154  
DB 237 MMLALAGPIFFLELSISGQFASQGVSWKAPLALOGGIMLLISVLAIYVVICYT 296  
QY 155 LYIAMSFOATLPNAICOPEN-----ENCYPSD-PLI-----AASVNI 192  
DB 297 LEYLFASFVSVLPMGSCNNPMTPECKDKTKLLDSCVIGHPKIQIKNSTFCMTAVPNTL 356  
QY 193 T--NGTSSA-----OLYFLRTVLOQSDIGEGGAPLWVLCFLTAMLVGVAVAR 242  
DB 357 TMAVNTSQANKTPVSGSEEFKYFVKISAGIEP-GEIRMPALCLFLAMVIVASLAG 415  
QY 243 GYKSSGKAAYFLALFPYVYMTLFTITIIIPGATDGLFEVTPQMAKLELGYVSAVQ 302  
DB 416 GIKTSKGVYVFAFPYVAVVILLIRGYTLPGAGAGIWTTPKMEKLDATVWKDAATQ 475  
QY 303 VFESLTVCTGPIIMSSYNGRHRNIYRDAMITVTLDTFTSLSGCTIFGILNLAELNS 362  
DB 476 IFFSLSAMGGILITLSSYNKFNHNCYRDTLVICTNSATSIFFAGVYFVSIGFMANERKV 535  
QY 363 EYGDVAVAGTSLAFISYPAIAKTQPOLSVLFFLMSVIGSGSVALLSTFNLTAM 422  
DB 536 NIENVADQ-GGCIATVVPALTRPLSPFWAITFFLMLTLTGIDTFMATETIYTSID 594  
QY 423 APPRVPTVYMSAMT--CSCGFLGLVYCTPGGOYILELVHDYGFELVFCALISELAGV 479  
DB 595 EEPKYLRTHKPVPFTLGCCICFFIMGPFMTQGGIYMFQVDTYASALVIAIFELVGI 654  
QY 480 FWIYGLNLCDIEEMLGKKTGAVWRLCGVITTPAINTVFFYALLASNNLVGDNYVTP 539  
DB 655 STVYGLORCEDIEMMIGFOPNIFWKVCMAFVPTLITFLICFSFYQWEMETYG-SYRPN 713  
QY 540 TAGVYSGYLMFLGTFVPIGIFSLKYRTGFSEETIKKAFKSPKSPGRSPRERE 597  
DB 714 NMSVYGLMMLACSVIMPIFMVYIKMH-LAPGRFIERKLKVCSPQDMPGLAQHGE 770

## RESULT 9

US-07-879-617A-8  
Sequence 8, Application US/07879617A  
Patent No. 5580775

GENERAL INFORMATION:  
APPLICANT: Tremean Jr., Robert T.  
APPLICANT: Caron, Marc G.  
APPLICANT: Blakely, Randy D.  
TITLE OF INVENTION: A High Affinity L-Proline Transporter  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 1100 Peachtree Street, Suite 2800  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: U.S.  
ZIP: 30309  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/879,617A

FILED DATE: 19920501  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: EMU109  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-815-6508  
TELEFAX: 404-815-6555  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 667 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Rattus  
TISSUE TYPE: Brain  
IMMEDIATE SOURCE:  
LIBRARY: rat forebrain cDNA library  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 76..95  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 103..127  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 147..167  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 247..266  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 272..294  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 321..341  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 353..375  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 404..427  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 454..473  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 487..509  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 530..549  
OTHER INFORMATION: /note= "Membrane-spanning domain"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 567..589  
OTHER INFORMATION: /note= "Membrane-spanning domain"

FEATURE:  
NAME/KEY: Region  
LOCATION: 44..45  
OTHER INFORMATION: /note= "protein kinase C phosphorylation site"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 71..72  
OTHER INFORMATION: /note= "cAMP-dependent protein kinase phosphorylation site"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 269..270  
OTHER INFORMATION: /note= "protein kinase C phosphorylation site"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 630..631  
OTHER INFORMATION: /note= "protein kinase C phosphorylation site"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 106..127  
OTHER INFORMATION: /note= "Leucine zipper motif"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 212..213  
OTHER INFORMATION: /note= "N-linked glycosylation site"  
US-07-879-617A-8

Query Match 30.7%; Score 1030.5; DB 1; Length 667;  
Best Local Similarity 38.7%; Pred. No. 8,3e-99;  
Matches 234; Conservative 94; Mismatches 255; Indels 21; Gaps 13;

QY 11 ESSEPKMEPKRSSQISL-PPANKKALDNTD--DTDLAE-PPERWWSNIEFLMSCIA 66  
DB 23 ESEQENCEMKKLOEAHLRKFPVPTDLMTSDQDQDVLDFADRNKWTGKDLDFLSCIG 82  
QY 67 TSVGLGNVWRFPIAYONGGAFVLPVYVILLVYGVKPPVYVEVLOFSSRNKVKWSTIS 126  
DB 83 YCVGLGNVWRFPIRAYTNGGAFVLPVYVILLVYGVKPPVYVEVLOFSSRNKVKWSTIS 142  
QY 127 PANKGTGYAQACCGYILSYVYVIGLCVYILAMSEFOATLPMAICPEW--ENCV-----P 180  
DB 143 PLFKGAGAMLLIVGLVAYIYNNITAYVLEFYELASLTSNLPWHCNMNTEGCLHRGP 202  
QY 181 SDPTLAAYNNITNGSSAQLVFLRYVL--QQSDGIEGIGAPVWVLYVLCFLPAMLMVFG 238  
DB 203 KDSNGALPL-NLSSTVSPSEEWWSRYVLHIQSGGI-GRGELRNWICLCLLAWIVFL 260  
QY 239 VVARGVKSSGKAAYFLALFPYVVMVITLTITLLPGATGILFVTPOMAKLLELGVWYS 298  
DB 261 CLKGVKSSGKAVYVFPATPRLILMLLVRCVTLPPAMKGIQVTLPPORHLLSSRWVIE 320  
QY 299 AVTVQVFSLTCTGPTLNFSSYNGERHNIYRDAMIYVTLDTFTSLSGCTIGILGNLAY 358  
DB 321 AALQIFYSLSGVGGGILLFASVYTFHQNIYDFVTLGNALTSILAGRAIFSVLGYMQ 380  
QY 359 ELNSEVGDVVGAGSTLAFISYDPAKAFQQLFVLFELMMSVYGISSVALLSTFTNT 418  
DB 381 ELGVPV-DOYAKAGPLAVIYIPQANTMLPLSPFWSLFFEMILITIGLDSQFAFLETIYT 439  
QY 419 LAMDAPP--RVPVYVMSAMTSCGFLGLVYCTPGQGYILLEVHYGGTFVFLCAISE 475  
DB 440 AVTDEPYYLRPKKAVFSGLLICVAMVIMLILITDGGMWVLLDDYSAPGLAMVYIT 499  
QY 476 LAGVNTITGLNCLDIEMLGKKGAVRWLCWGYITPAIMKTVFFYALLASNNLVFGN 535  
DB 500 CLAVTVVVGIOFRCRDIHMLGFKPGLYFRACMLFSPATLLALVYSIVKYQPSBYG-S 558  
QY 536 YVPTAGVSGVLMFLGNTFVPIGIGFSLYKRYGTSETIKKAFHSPSGPSPRRR 595

Db 559 YRRPAPAEILGILMGILSCLMIPACMLAVLR-EEGSLMERLQQAAPRAIDMGF-SLEEN 616

QY 596 REMM 599

Db 617 RTGM 620

RESULT 10  
 US-08-753-985-8  
 Sequence 8, Application US/08753985  
 Patent No. 5759788  
 GENERAL INFORMATION:  
 APPLICANT: Fremeanu Jr., Robert T.  
 APPLICANT: Caron, Marc G.  
 APPLICANT: Blakely, Randy D.  
 TITLE OF INVENTION: A High Affinity L-Proline Transporter  
 NUMBER OF SEQUENCES: 13  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Kilpatrick & Cody  
 STREET: 1100 Peachtree Street, Suite 2800  
 CITY: Atlanta  
 STATE: Georgia  
 COUNTRY: U.S.  
 ZIP: 30309  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/753,985  
 FILING DATE: 03-Dec-1996  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/879617  
 FILING DATE: 01-MAY-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Padst, Patrea L.  
 REGISTRATION NUMBER: 31,284  
 REFERENCE/DOCKET NUMBER: EMU109  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 404-815-6508  
 TELEFAX: 404-815-6555  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 667 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: N-terminal  
 ORIGINAL SOURCE:  
 ORGANISM: Rattus  
 TISSUE TYPE: Brain  
 IMMEDIATE SOURCE:  
 LIBRARY: rat forebrain cdna library  
 CLONE: rTB2-20  
 FEATURE:  
 NAME/KEY: Domain  
 LOCATION: 76..95  
 OTHER INFORMATION: /note= "Membrane-spanning domain"  
 FEATURE:  
 NAME/KEY: Domain  
 LOCATION: 103..127  
 OTHER INFORMATION: /note= "Membrane-spanning domain"  
 FEATURE:  
 NAME/KEY: Domain  
 LOCATION: 147..167  
 OTHER INFORMATION: /note= "Membrane-spanning domain"  
 FEATURE:  
 NAME/KEY: Domain

```

LOCATION: 247..266 /note= "Membrane-spanning domain"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 272..294
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 321..341
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 353..375
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 404..427
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 454..473
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 487..509
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 530..549
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 567..589
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Region
LOCATION: 71..72
OTHER INFORMATION: /note= "protein kinase C
phosphorylation site"
FEATURE:
NAME/KEY: Region
LOCATION: 269..270
OTHER INFORMATION: /note= "protein kinase C
phosphorylation site"
FEATURE:
NAME/KEY: Region
LOCATION: 630..631
OTHER INFORMATION: /note= "protein kinase C
phosphorylation site"
FEATURE:
NAME/KEY: Region
LOCATION: 106..127
OTHER INFORMATION: /note= "Leucine zipper motif"
FEATURE:
NAME/KEY: Region
LOCATION: 212..213
OTHER INFORMATION: /note= "N-linked glycosylation
site"

```

	Query Match	30.7%	Score 1030.5	DB 1	Length 667
	Best Local Similarity	38.7%	Pred. No. 83e-89		
	Matches 234	Conservative 94	Mismatches 255	Indels 21	Gaps 13
QY	11 ESSEPKMKRSSQSL-PPANKKALDNID--PTDLEAF-PPERWMSNNIEFLMCSIA	66			
DB	23 ESEQNCMKKIQEHLHNRKPTVPTDLMTPSOGDDVDVDAADGNGNTGKIDFLSIG	82			

QY 67 TSVGLGNVRRPFLAYONGGAFPLPYVYVLLLVGKPYVYLVGQFSSHNKYSIS 126  
 Db 83 YCVGLGNVRRPFLAYONGGAFPLPYVYVLLLVGKPYVYLVGQFSSHNKYSIS 142  
 QY 127 PAMGCTGAAAGCGILSYVWTCGCLYTLANSFOATLPAICQPM--ENCV---P 180  
 Db 143 PLEFGAAGAMLLIYGLVAYVMTAVLEFLFASLTSNLPWEHGMWMTRELERHGP 202  
 QY 181 SDPLASVNNITNGTSSAQLFPLRVL--QOSDIGGAGAPIMYLVLCFLIMLVFG 238  
 Db 203 KDGNGALPL-NLSTVSPSEKYSRYVLIHQSGGT-GRGEIWMNLCLLMLWVIVFL 260  
 QY 239 VVARGVSSGKAAYELALFPYVNMETLFTITLLPGATDGLFEPVQMAVLELGYWS 298  
 Db 261 CILKGVKSSGKVYVFTATFPLTLMLLVKGVTLPGAMKGIQVFLTFQFHLSSKVWIE 320  
 QY 299 AVTQVFEFLVTCGPIIMFSSYNGERNIRYRDAMIVTLDITFSLGCTIFGLGNLAY 358  
 Db 321 AALQIFSLGVGEGGLTFEASYNTHONIRDFIVTLGNATISILAGFALFSLGYMSQ 380  
 QY 359 ELNSEVDVYGAGCTSLAFTSYPAIAKTFQPOLFSVLFLMSVLGSSVALLSTENT 418  
 Db 381 ELGVPV-DQYAKAGPGLAFVYIPQAMTMLPSFPWSLFFPMILTLGDSQFALLETIVT 439  
 QY 419 LADAPF--RVPLVYNSAMTCSGIFLLGLVYCTPGQVILLEVYHGGFFVLFCALSE 475  
 Db 440 AVTDEFYVLRPKKAVFSGLICVAVYLMGLILTTDGMVYLVLLDDYSASGLWVVYIT 499  
 QY 476 LAGVFWYGLNCLDIEFIMGKKGAYWRLCWGVTPAINTVFFALLASNNLVGDN 535  
 Db 500 CLAVTRYGIQRCDRIHMLGFRGLEYFACWLELSPALLALVSYVYQSEVG-S 558  
 QY 536 YVYPTAGVSGYIMLFGMTFVPLSGFSLYKRYGTGTFSETIKAFKSKSMGPRSRER 595  
 Db 559 YRPPAMAILGILMGLSLCMLIPASMLVAVLR-EGSLMRLQASPAIDWGP-SLEEN 616  
 QY 596 REMM 599  
 Db 617 RTGM 620

TELEFAX: 404-815-6555  
 INFORMATION FOR SEQ ID NO: 9:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 635 amino acids  
 TYPE: AMINO ACID  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: Protein  
 HYPOTHEICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: N-terminal  
 ORIGINAL SOURCE:  
 ORGANISM: Rattus  
 TISSUE TYPE: Brain - Proline Transporter  
 FEATURE:  
 NAME/KEY: Domain  
 LOCATION: 46..65  
 OTHER INFORMATION: /note= "Proposed transmembrane domain."  
 FEATURE:  
 NAME/KEY: Domain  
 LOCATION: 72..97  
 OTHER INFORMATION: /note= "Proposed transmembrane domain."  
 FEATURE:  
 NAME/KEY: Domain  
 LOCATION: 117..137  
 OTHER INFORMATION: /note= "Proposed transmembrane domain."  
 FEATURE:  
 NAME/KEY: Domain  
 LOCATION: 217..236  
 OTHER INFORMATION: /note= "Proposed transmembrane domain."  
 FEATURE:  
 NAME/KEY: Domain  
 LOCATION: 243..264  
 OTHER INFORMATION: /note= "Proposed transmembrane domain."  
 FEATURE:  
 NAME/KEY: Domain  
 LOCATION: 291..311  
 OTHER INFORMATION: /note= "Proposed transmembrane domain."  
 FEATURE:  
 NAME/KEY: Domain  
 LOCATION: 322..345  
 OTHER INFORMATION: /note= "Proposed transmembrane domain."  
 FEATURE:  
 NAME/KEY: Domain  
 LOCATION: 373..397  
 OTHER INFORMATION: /note= "Proposed transmembrane domain."  
 FEATURE:  
 NAME/KEY: Domain  
 LOCATION: 424..443  
 OTHER INFORMATION: /note= "Proposed transmembrane domain."  
 FEATURE:  
 NAME/KEY: Domain  
 LOCATION: 456..479  
 OTHER INFORMATION: /note= "Proposed transmembrane domain."  
 FEATURE:  
 NAME/KEY: Domain  
 LOCATION: 500..519  
 OTHER INFORMATION: /note= "Proposed transmembrane domain."  
 FEATURE:  
 NAME/KEY: Domain  
 LOCATION: 536..559  
 OTHER INFORMATION: /note= "Proposed transmembrane domain."

RESULT 11  
 US-07-879-617A-9  
 ; Sequence 9, Application US/07879617A  
 ; Patent No. 5580775  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fremieu Jr., Robert T  
 ; APPLICANT: Caron, Marc G  
 ; APPLICANT: Blakely, Randy D  
 ; TITLE OF INVENTION: A High Affinity L-Proline Transporter  
 ; NUMBER OF SEQUENCES: 13  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Kilpatrick & Cody  
 ; STREET: 1100 Peachtree Street, Suite 2800  
 ; CITY: Atlanta  
 ; STATE: Georgia  
 ; COUNTRY: U.S.  
 ; ZIP: 30309  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/879,617A  
 ; FILING DATE: 19920501  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Pabst, Patrea L.  
 ; REGISTRATION NUMBER: 31,284  
 ; REFERENCE/DOCKET NUMBER: EMU03  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 404-815-6508



OTHER INFORMATION: domain."  
FEATURE:  
NAME/KEY: Active-site  
LOCATION: 76..77  
OTHER INFORMATION: /note= "Leucine zipper motif"  
FEATURE:  
NAME/KEY: Active-site  
LOCATION: 83..84  
OTHER INFORMATION: /note= "Leucine zipper motif"  
FEATURE:  
NAME/KEY: Active-site  
LOCATION: 90..91  
OTHER INFORMATION: /note= "Leucine zipper motif"  
FEATURE:  
NAME/KEY: Active-site  
LOCATION: 97..98  
OTHER INFORMATION: /note= "Leucine zipper motif"  
US-07-879-617A-9

Query Match 30.3%; Score 1017; DB 1; Length 635;  
Best Local Similarity 39.4%; Pred. No. 1.5e-87;  
Matches 225; Conservative 91; Mismatches 237; Indels 18; Gaps 11;

QY 41 DTDLAE-PPERWYNNIEFLMSCIATSVGLNVMRPFYAYONGGAFIYVYVLL 99  
D 26 DVDLDVDFADRGWMTKIDFLSLCIGYCVGLGNVWRFYRAYNGGAFIYVYVLL 85  
QY 100 VGRFVYIECVLGGFSSRNKYSKVPAMKGTGACAGGILSYVYVCGICLYYLA 159  
D 86 CGIPLEFLSLGOFSSGLPIAWKISPLFEGAGAMLLYGLAITYNMIAVLYL 145  
QY 160 MSFOATLPMAICPEM--ENCV---PSDPLASVNNITNGTSSADLYFRLV--QOS 211  
D 146 ASLYSNLPWEHCGWMTTERCLEHRGPKDNGALPL-NLSTVSPSEVYRWYVLIHGS 204  
QY 212 DGEIGGAPWYVYVLCFLTAMLVGVNAGVSSGKAYFLLFPYVMTITFTTII 271  
D 205 QGI-GRGELRMNLCILLLAWIVFLCIIRKGVSSGRVYFTATFYLILMLLVGVT 263  
QY 272 LPGATDGLFEFVPOMAKLELGVYSANVQVFSLVNCTGPIFMSSYNGFRNIRDA 331  
D 264 LPGAOKIOYLTQFHLHLSKAYIMEAADIYSLGVGEGGLTFPSYNTFHONIRDT 323  
QY 332 WIVTTLDTFSLGCTIFGLNMAELNSEVGVAGGTSIAFTSYDDAIKTRPOQ 391  
D 324 FIYTLGNATSIILAGFAIFSVLGYSGDELGPV-DYAKAGPGLAFIYPPAMMLPLSP 382  
QY 392 LFSVLEFLMSVIGTSSVALLSTFTNLADAPP--RVPTVYSAMTSCSGFILGLVYC 448  
D 383 FWSFLFEMLLTGLDQFALEFIVAVTDEFPYLRPKKAVFSGLICVAMYLMLLT 442  
QY 449 TPGGOYILELVHDHGFVLFCAISELAGFMYIGLENCLDIEFMIGKTGAYNRLCW 508  
D 443 TDGGMVYLLDDYSASFGMLVYVITCLAVTRYGIQRDRDHHMLGEPGLYFRACW 502  
QY 509 GVTPAIVMTVEFYALLASNNLVFGDNYVPTAGVSGYLMFLGTFVPIGIFSXYKY 568  
D 503 LFLSPATILALLVYSIKYQPSYRG-SYRPPAMAEILGILMGLSCIMIPAGMLVAVLR- 560  
QY 569 RTGFESETIKKAFKSPKSPGRPREREM 599  
D 561 EESGLMERLQOASRPADWGP-SLEENRTGN 590

RESULT 12  
US-08-753-985-9  
Sequence 9, Application US/08753985  
Patent No. 5759788  
GENERAL INFORMATION:  
APPLICANT: Fireneau Jr., Robert T.  
APPLICANT: Caron, Marc G.  
APPLICANT: Blakely, Randy D.  
TITLE OF INVENTION: A High Affinity L-Proline Transporter

NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 1100 Peachtree Street, Suite 2800  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: U.S.  
ZIP: 30309  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/753,985  
FILING DATE: 03-DEC-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/879617  
FILING DATE: 01-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Padst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: EMU109  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-815-6508  
TELEFAX: 404-815-6555  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 635 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Rattus  
TISSUE TYPE: Brain - Proline Transporter  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 46..65  
OTHER INFORMATION: /note= "Proposed transmembrane  
OTHER INFORMATION: domain."  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 72..97  
OTHER INFORMATION: /note= "Proposed transmembrane  
OTHER INFORMATION: domain."  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 117..137  
OTHER INFORMATION: /note= "Proposed transmembrane  
OTHER INFORMATION: domain."  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 217..236  
OTHER INFORMATION: /note= "Proposed transmembrane  
OTHER INFORMATION: domain."  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 243..264  
OTHER INFORMATION: /note= "Proposed transmembrane  
OTHER INFORMATION: domain."  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 291..311  
OTHER INFORMATION: /note= "Proposed transmembrane  
OTHER INFORMATION: domain."  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 322..345

Query Match	30.3%	Score	1017	DB	1	Length	635
Best Local Similarity	39.4%	Pred	No. 1.5e	87			
Matches	225	Conservative	91	Mismatches	237	Indels	18
						Gaps	11

```

Yy 392 LEVSLFLPLMSVLDIGSSVALLSPENTLAMAF---RVPIVYMSAMKSCGFLGIYVC 448
Db 383 FWSLFFPEMLLTLDIGDSQFALETTVIAVTEEPFYLPKRAVESGLICVANYLMLGLIT 442
Yy 449 TPGQYLIELVDHNGGTFVLVFCASISLAGVFIYGLLENLCIDIEFMIGKTKGAYWRLCW 508
Db 443 TDGGMVNYVLDDYASASGRLMVVYITGLAVTRYGIORPCRDIMMIGFPGULYFRACW 502
Yy 509 GVTPALMTVTFYFFALLASNNLVGDNVYVTPAGVSYLMLFLGTFVPIPIGISLKY 568
Db 503 LFLSPATLALLVYSIVYQPSYEG-SYRFPAMELLGILMLLSCMLIPAGMLVAVLR- 560
Yy 569 RTGFSEFTIKKAFSKPSMGCRSPREREXEM 599
Db 561 EEGSLMERLQOASRAIDMGP-SLEENRTGM 590

RESULT 13
US-08-301-722A-5
Sequence 5, Application US/08301722A
Patent No. 5756307
GENERAL INFORMATION:
APPLICANT: Uhl, George R.
APPLICANT: Vandenberg, David
APPLICANT: Persico, Antonio
TITLE OF INVENTION: SEQUENCE OF HUMAN DOPAMINE TRANSPORTER
TITLE OF INVENTION: CDNA
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESS: Birch, Stewart, Kolasch & Birch, LLP
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-3487
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/301,722A
FILING DATE: 07-SEP-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1173-406P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 599 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: Protein
LOCATION: 1..599
OTHER INFORMATION: /note= "Hgbat sequence, see Fig. 5"
OTHER INFORMATION: 5"
US-08-301-722A-5

Query Match 30.1%; Score 1011; DB 1; Length 599;
Best Local Similarity 35.4%; Pred. No. 4.9e-87;
Matches 213; Conservative 118; Mismatches 229; Indels 42; Gaps 13;

```

Db 10 DGOI-----STEVSAPYANDKPKTLVYKQKADLPDRDTWKGRDFEL 54  
Qy 62 MSCATSVGLGNWMPFPIAYONGGAFLLVPYVILLVGRPVYLLGCVLGQFSSRNSVK 121  
Db 55 MSCGYAIGLGNWMPFPIAYONGGAFLLVPYVILLVGRPVYLLGCVLGQFSSRNSVK 114  
Qy 122 VMSISPAKGTGYAQAACGYILSYVYVIGCLTYLAMSFOATLPMAICQPEW--ENCY 179  
Db 115 VMKLAPMKGYGLAAVLSFNLNTIYIYISMAIYLYLNSFTTLPMKQCDNPMWTDKCF 174  
Qy 180 PSDPTLAASVNNITNGTSSAQLYFRLVLOOSDIEGGLGAPWYLVLCFLPAMLVPGV 239  
Db 175 SN-----YSWNTMTMSAVVEFERNMHQMTDIDKPG-QDIRMPLATTLAIWLVYFC 228  
Qy 240 VARGVSSGKAAYFLATLPYVMTLFTTTLLPGATGGLIFEVTPQMAKLLGQWYSA 299  
Db 229 IMKGVMTGKVVYSATVPYIMLLIFPRGYTLPEAKGILFYITPNRKLSDSEVWIDA 288  
Qy 300 VTQVFFSLTVCPTIMESSYNGFRHNYRDAMVITTLDTFTSLSGCTIGILGNLAVE 359  
Db 289 ATQIFFSYGLIGSLIALGYSNPFHNNYRDSIIVCCINSCTSMFAGVIFSIYGFMAHV 348  
Qy 360 LNSEGVYVAGGTSIAFISYPAIDAKTFQPOLPSVLEFLMVSIGSSVALLSTFNTL 419  
Db 349 TKRSIAD-VAASGGLAFPLAEVATQPLISPLMALIFFSMLMIGDISOCTYEGFTTA 407  
Qy 420 AMDAPRY-----PYVMSAMTSCGFLGLGYCTPGGOYILEVDHYGFTL-VLFCAS 474  
Db 408 LVDEYPRLLRNRRELFLAAV-CIISYLLGLSNITOGGIYVFKLPFYASASGMSLFLVFF 466  
Qy 475 ELAGVFWYGLNCLDIEFMLGKKTGYWRLCWGVTPAINTVFEFALLASNNLVFGD 534  
Db 467 ECVSISWYGYNRFPDNIQENWVGRPCIMWKLCSFEPPIIYAGVIFSAVOMPLTNG- 525  
Qy 535 NYVYPTAGYVGYLMLFLGMTFVPIGIFSLYKRYT--GTSEPTIKKAFHS-----KRW 587  
Db 526 NYVEPKMGQGVGMALSSMVLIP---GYMAVMTLALGSLKORLOVWQPSDEDTVRBN 582  
Qy 588 GP 589  
Db 583 GP 584

RESULT 14  
US-08-295-814E-11  
Sequence 11, Application US/08295814E  
Patent No. 5658786  
GENERAL INFORMATION:  
APPLICANT: Smith, Kelli E.  
APPLICANT: Borden, Laurence A.  
APPLICANT: Hartig, Paul R.  
APPLICANT: Weinschank, Richard L.  
TITLE OF INVENTION: DNA ENCODING TAURINE AND GABA  
TITLE OF INVENTION: TRANSPORTERS AND USES THEREOF  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/295,814E  
FILING DATE: DECEMBER 19, 1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John

REGISTRATION NUMBER: 28, 678  
REFERENCE/DOCKET NUMBER: 40558-B-PCT-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 599 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-295-814E-11

Query Match 30.1%; Score 1008; DB 1; Length 599;  
Best Local Similarity 34.8%; Pred. No. 9,5e-87;  
Matches 211; Conservative 122; Mismatches 231; Indels 42; Gaps 13;

Qy 3 DGOVNGGFESSEPKMKPRSSQIS-LPPANNKALDNIDPDLEAPPERMWSNNIEFL 61  
Db 10 DGOI-----STEVSAPYANDKPKTLVYKQKADLPDRDTWKGRDFEL 54  
Qy 62 MSCATSVGLGNWMPFPIAYONGGAFLLVPYVILLVGRPVYLLGCVLGQFSSRNSVK 121  
Db 55 MSCGYAIGLGNWMPFPIAYONGGAFLLVPYVILLVGRPVYLLGCVLGQFSSRNSVK 114  
Qy 122 VMSISPAKGTGYAQAACGYILSYVYVIGCLTYLAMSFOATLPMAICQPEW--ENCY 179  
Db 115 VMKLAPMKGYGLAAVLSFNLNTIYIYISMAIYLYLNSFTTLPMKQCDNPMWTDKCF 174  
Qy 180 PSDPTLAASVNNITNGTSSAQLYFRLVLOOSDIEGGLGAPWYLVLCFLPAMLVPGV 239  
Db 175 SN-----YSWNTMTMSAVVEFERNMHQMTDIDKPG-QDIRMPLATTLAIWLVYFC 228  
Qy 240 VARGVSSGKAAYFLATLPYVMTLFTTTLLPGATGGLIFEVTPQMAKLLGQWYSA 299  
Db 229 IMKGVMTGKVVYSATVPYIMLLIFPRGYTLPEAKGILFYITPNRKLSDSEVWIDA 288  
Qy 289 ATQIFFSYGLIGSLIALGYSNPFHNNYRDSIIVCCINSCTSMFAGVIFSIYGFMAHV 348  
Qy 300 VTQVFFSLTVCPTIMESSYNGFRHNYRDAMVITTLDTFTSLSGCTIGILGNLAVE 359  
Db 289 ATQIFFSYGLIGSLIALGYSNPFHNNYRDSIIVCCINSCTSMFAGVIFSIYGFMAHV 348  
Qy 360 LNSEGVYVAGGTSIAFISYPAIDAKTFQPOLPSVLEFLMVSIGSSVALLSTFNTL 419  
Db 349 TKRSIAD-VAASGGLAFPLAEVATQPLISPLMALIFFSMLMIGDISOCTYEGFTTA 407  
Qy 420 AMDAPRY-----PYVMSAMTSCGFLGLGYCTPGGOYILEVDHYGFTL-VLFCAS 474  
Db 408 LVDEYPRLLRNRRELFLAAV-CIISYLLGLSNITOGGIYVFKLPFYASASGMSLFLVFF 466  
Qy 475 ELAGVFWYGLNCLDIEFMLGKKTGYWRLCWGVTPAINTVFEFALLASNNLVFGD 534  
Db 467 ECVSISWYGYNRFPDNIQENWVGRPCIMWKLCSFEPPIIYAGVIFSAVOMPLTNG- 525  
Qy 535 NYVYPTAGYVGYLMLFLGMTFVPIGIFSLYKRYT--GTSEPTIKKAFHS-----KRW 587  
Db 526 NYVEPKMGQGVGMALSSMVLIP---GYMAVMTLALGSLKORLOVWQPSDEDTVRBN 582  
Qy 588 GPRSR 593  
Db 583 GPEOPQ 588

RESULT 15  
US-08-240-783B-4  
Sequence 4, Application US/08240783B  
Patent No. 5756348  
GENERAL INFORMATION:  
APPLICANT: Smith, Kelli  
APPLICANT: Borden, Laurence A.  
APPLICANT: Hartig, Paul R.  
APPLICANT: Weinschank, Richard L.

TITLE OF INVENTION: DNA ENCODING A GLYCINE TRANSPORTER AND USES  
 TITLE OF INVENTION: THEREOF  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Cooper & Dunham LLP  
 STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.14  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/240,783B  
 FILING DATE:  
 CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: White, John P.  
 REGISTRATION NUMBER: 28,678  
 REFERENCE/DOCKET NUMBER: 1794/39875-A-PCT-US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 278-0400  
 TELEFAX: (212) 391-0525  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 599 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHEICAL: N  
 FRAGMENT TYPE: Internal  
 ORIGINAL SOURCE:  
 ORGANISM: RAT GABA TRANSPORTER (GAT-1)  
 US-08-240-783B-4

Query Match 30.1%; Score 1008; DB 1; Length 599;

Best Local Similarity 34.8%; Prd No. 9.5e-87;  
 Matches 211; Conservative 122; Mismatches 231; Indels 42; Gaps 13;

QY 3 DGOVNGFESSEPKMEPKRSSQIS LPPANKKALDNIDDTLEAPPERMVMNNIEFL 61  
 DB 10 DGOI-----STENSHAPASDKPKTLVVKQKAGDLPDRDTWKGRDFL 54  
 QY 62 MSCATSVGLGNWRFPIAYONGGAFVPIYVILLVGKPYVLECVLGQSSRNSVK 121  
 DB 55 MSCGVYIGLGNWRFPIYLCGKGGAFLLPYFTLLFAGVPLLECSLGQYTSIGLG 114  
 QY 122 VMSISPAKGTGYAQAAGCYILSYVYVIGLCLYLAMSFOATLPWAICPEW--ENCV 179  
 DB 115 VMKLAPEKGVGLAANLFWLMIYIVLISMALYLYNSFTTLLPWKQCDNPWNTDRCF 174  
 QY 180 PSDFTLAASVNNITNGTSSAQLFTRTVLOQSDGIEGGLGAPIWLVLCFIAMLVFV 239  
 DB 175 SN-----YSLVNTNMTSAVVEWIRNMHOMTDGLDRP-QQIRWPLAITLAIAVLYFC 228  
 QY 240 VARGVSSGKAAYFLAEPVVMITLFTITIIIPGATDGLFFTPQMAKLLIEGVYXA 299  
 DB 229 IMGVGWTKGVVFSATYPIIMILTFEFGVTLPGAKEGILFYTPNFRKLSDEVWLDA 288  
 QY 300 VTOVFESLAVCTGPIIMFSSYNFHNHYDAMIVTLDFTFSLSCCTFGILGNLAYE 359  
 DB 289 ATQIFEFYGLGLSLILASYNFNNVYDSDIIVCCINCSIMFAGFVIFSTVGMAYH 348  
 QY 360 LNSEVGVAGAGTSLAFISYPIIAKTFQPLFVFLFMSVYLGIGSSVALLSTPNTL 419  
 DB 349 TKRSIAD-VAASGGLAFIAYPAVATQLPSPLAILFESMLMLGLIDSOFCVEGFITA 407  
 QY 420 AMDAFPRV-----PIVYMSAMTCGFLGLGVYCTPGQYILLEVHDHGGTFLL-VLPCAIS 474  
 DB 408 LVDEYPRLLRNRLRLTAAY-CIVSYLIGLSNITGGIYVFKLEDYYSASGMSLLFLVFF 466

QY 475 ELAVGEWVIGLENICLDIEEMIGKKGAVYRLQMGVITPAIMTVEFYALLASNNLYFGD 534  
 DB 467 ECYSISWFYGVNRRPYDNIQEMVGRPCITWKKLMSFPTPLIVAGVFLESVAVOMTPLTMG- 525  
 QY 535 NYVYPTAGVYGYLMLFLGNTFVPIGIFSLYKRT--GTFSETIKKAFHS-----KPSW 587  
 DB 526 SYVEPKWGGGVGMMLALSMVLLP---GYMAYMFLTKSLKQRLQVMIQPSEDIYVRPN 582  
 QY 588 GPRSPR 593  
 DB 583 GPEQPQ 588

Search completed: July 21, 2003, 09:28:09  
 Job time : 20 secs

[illegible]

QY 174 EM-----ENCPSU-----AASVNNIR--NGTSSA-----QL 201  
DB 314 PNNTPCKDKTKLLDSDSCVISHKIQIKNSHFCMTAIPNVTMNFSTOANKTEVSGSEE 373  
QY 202 YELRVYLOOSDIEGGLGAPVYVYVLCLEFIAMLVFVGAVRGSKSGRAAYFLAPYV 261  
DB 374 YKRYVYKLSAGIEVP-GEIRMPLELCLEFLAVVYVYASIAKIKTSKGVYVETAFEPYV 432  
QY 262 MTLFPTTILPGADGILFEVTEPMAKLLEGVYKSAVVOVFESLVCYCTGPIIMFSSYN 321  
DB 433 LVILLIRGYTLPGAAGIWTFTETPKMEKLTATVAKDQATQIFESLSAMGGLTLLSSN 492  
QY 322 GERNNIYNDAMIVTTLDTFTSFLSCCTIFGILGNLVELNSEVDGVAGGTSIAFISYP 381  
DB 493 KFNHNCYRDLTITCTNSATSHFAFVIFSVIGFMANRKNIEVADQ-GRGIAFVYVP 551  
QY 382 DALIAFTQDQFVSVEFLMMSYLGSSVALLSTFNTLAMPAPRPVTVYSAMT---CS 438  
DB 552 EALTPLPSFPWALFEFLMLTLGLDTMFATLETIVISISDEFPKYLTHKPVFTLGCCI 611  
QY 439 CGFLGLVYCTPGQYILELVNDHYGTFVLEFCALISELAGVEMYLENLCIDIEPMUGK 498  
DB 612 CFEIFGPMITOGGIMQVLDVYNASALVITIAIFELVGIYVGLQRCEDIEEMMIGF 671  
QY 499 KTGAVWRKCGVITPAITVTFYVLLASNNLVFGDNVYPTAGVSGYLMFLGTFVP 558  
DB 672 QPNIFWKVMAFVPTLITFILCFSEYQMEBMTYG-SYRYPNMSVGLMLLACSVMIP 730  
QY 559 IGIGFSLKYKRTGFSETIRKAFHFKPSMGPRSRERE 597  
DB 731 IMFVTKMH-LAPGRFIERKLAESCQDMGPFLLAQHGE 768

RESULT 2  
US-09-795-693-24

Sequence 24, Application US/09795693  
Patent No. US20020068710A1  
GENERAL INFORMATION:  
APPLICANT: Glucksmann, Maria A.  
TITLE OF INVENTION: 20685, 579, 7114, 23821, 33894, and  
TITLE OF INVENTION: 32613, No. US20020068710A1 Human Transporters  
FILE REFERENCE: 35800/209292  
CURRENT APPLICATION NUMBER: US/09/715,693  
CURRENT FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: 60/185,906  
PRIOR FILING DATE: 2000-02-29  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 24  
LENGTH: 624  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Pfam consensus sequence  
US-09-795-693-24

Query Match 30.6%; Score 1025.5; DB 10; Length 624;  
Best Local Similarity 37.6%; Pred. No. 1.7e-84;  
Matches 234; Conservative 94; Mismatches 201; Indels 93; Gaps 15;

QY 51 RMWYNNIEFLMSCIATSVGLNVMRFPFIAYONGGAFVPIYVILLVGRKPVYLEGV 110  
DB 1 RETWMSKLDLFDVLSVGFVAVGLNVMRFPFLCYKNGGALLIPILIFLIYAGIDPLELA 60  
QY 111 LGGFSSRNSVYK-----STPAMGTGYQAAGCGYLLSYVVICGLCYLLAMS 161  
DB 61 LGGYTRGSSITVYRKRIIDKKGICELFGIGVASTVIAFYIGIYVNIAMALYLLFSS 120  
QY 162 FQATLPALCOPEME--NCV-----PSDPTLAA-SVNNIIN---GSSSQKFLKRV 207  
DB 121 FTTELPMATCNMSWNPNCVEREARENSINGSLAALSNLTDYLERKSPVEEERWERY 180  
QY 208 LQ--QSDGIEGGLGAPVYVYVLCLEFIAMLVFVGAVRGVKS-SGRAAYFLAPYVMT 264

DB 181 LKLESSGIE-DLGEIRWELTCLLLAMIVYFCLMKVSGSGKVVYFATPFYVLLV 239  
QY 265 LEFTIILPGATDGLFEFVPQMAKLELGWYSAVVOVFESLVCYCTGPIIMSSYNGFR 324  
DB 240 LILINGVILPGAADIKRYLPDPFSKLLDPQVMDAAVQIFESLIGIGVILALASYNKFH 299  
QY 325 HNIRDAMIYVTLDTFTSFLSGCTIFGILGNLA----- 357  
DB 300 NNCYRDAIIVAFINSITSFLAGFVIFSLGPMANIVQEGVPENEKILLSVSLRDLPH 359  
QY 358 -----VELNSEVD---VVC-----AGTSIAFTSYDAIAKTFQPO 391  
DB 360 VNLSALPADYSVDVIVISEVAESEFVLGLACLEDLKDVKQAGPGIAFYAPEATVPLSP 419  
QY 392 LFSVLPFLMSVILGSSVALLSTFNTLAMPAPRPVTVYSAMTSCGFLGIYV 447  
DB 420 FMAVLFPLMLTLGLDQFGGVEGIIATVDEPILIRKVRRELFILVCVIFSLGLEW 479  
QY 448 CTPEGQYILELVNDHYGTFVLEFCALISELAGVEMYLENLCIDIEFMUGKKTGAYWRL 506  
DB 480 VTEGCIYVFTLFPYVYASGFSLLFVVEFCIAVAMVYGDIDRFYDDITEMLGRPGLYMKL 539  
QY 507 CMGVTTPAINTVFEVLA-----LASNNLV-FGDNVYPTAGVSGYLMFLGTFVP 558  
DB 540 CMKFSVPLILFLEFISVYQYGLKPLTYNNMIREADYVPMANALGMLLASSMLCVP 599  
QY 559 IGIGFSLKYKRTGFSETIRKAFHFKPSMGPRSRERE 597  
DB 600 LVIIYKLLSTEGDSLIERLOKA 621

RESULT 3  
US-10-156-239-24

Sequence 24, Application US/10156239  
Publication No. US20030036074A1  
GENERAL INFORMATION:  
APPLICANT: Glucksmann, Maria A.  
TITLE OF INVENTION: No. US20030036074A1 Nucleic Acid Sequences Encoding Human Tr  
TITLE OF INVENTION: Atpase Molecule, A Human Ubiquitin Hydrolyase-Like Molecule, A  
FILE REFERENCE: 35800/247645  
CURRENT APPLICATION NUMBER: US/10/156,239  
CURRENT FILING DATE: 2002-05-24  
PRIOR APPLICATION NUMBER: 09/795,693  
PRIOR FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: 60/185,906  
PRIOR FILING DATE: 2000-02-29  
PRIOR APPLICATION NUMBER: 09/809,557  
PRIOR FILING DATE: 2001-03-15  
PRIOR APPLICATION NUMBER: 60/192,018  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: 01/808,568  
PRIOR FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 60/191,790  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: 09/808,767  
PRIOR FILING DATE: 2001-03-15  
PRIOR APPLICATION NUMBER: 60/191,781  
PRIOR FILING DATE: 2000-03-24  
NUMBER OF SEQ ID NOS: 60  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 24  
LENGTH: 624  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Pfam consensus sequence  
US-10-156-239-24

Query Match 30.6%; Score 1025.5; DB 15; Length 624;  
Best Local Similarity 37.6%; Pred. No. 1.7e-84;

Matches 234; Conservative 94; Mismatches 201; Indels 93; Gaps 15;

```
QY 51 RAVWSNNIEFLMCIATSVGLGNVWRPFPIAYONGGAFVLYVILLGKPYVYECV 110
    1 RETWSGKLDPLVSVGFVAVGLGNVWRPFPIAYONGGAFVLYVILLGKPYVYECV 60
QY 111 LCOFSSRNSVKW-----SISPAKGTGYAQAAGCGYILSYVYVIGCLCYLAMS 161
    61 LQYTRREGSITVWRKKILDKGKICPLFGIGIYASIVIAFYIGIYVNIAMLYLFSS 120
QY 162 FOATLPAICOPEME--NCV-----PSDPTLAA-SYNNITN---GTSSAQLYFLTV 207
    121 FTTELPMATCNNSMTNPNCEVEREAEINSTNSLAALSKNLTDTLERTSPVEEWEKGV 180
QY 208 LQ--OSDGLIEGGLAPIWYVLCFLIAMIWFGVAVGKVS-SGKAAYFLALFPYVMIT 264
    181 LKLSSESGIE-DIGELRWELTCLLAMIWYVFCIMKVGKSGSVVYFTAFPPYVILV 239
QY 265 LEFTTILPGATDGLFEVTPQMAKLELGVWYSANTOVFSLVCTGPTIMESSYNGFR 324
    240 LLIRGVTLPGADGKIFLTLPDSKLLDPQWIDAAQTIFSLGIGFGLIALASYNKFH 299
QY 325 HNIYRDAMIWTTLDFTSFLSGCTIFGLIGNLA----- 357
    300 NNCYRDALIVSFINSITSLAGFVIFSLIGFMANTVOEGVPENEKILLSVLSRDLPH 359
QY 358 -----YEINSEVGD---VVG-----AGTSLAFISYPAIAKTPQPO 391
    360 VNLSTADYDVYDISEVAESEFVLGACLEDELKVOAGPLAFIAPVAVMMLPSP 419
QY 392 LESVLEFLMMSVILGIGSSVALLSTFNLTAMDAFP-----RVPYVMSANTSCGGLLGVY 447
    420 FMAVLEFLMLTLTGIDSGOGVEGIIITLALVDEFPILLRKVRRELFILVCAISFLGLFEM 479
QY 448 CTGPGQYVILEVDHYGTF-FLVFCALSELGAVMIYGLIENICLIDIEFMGKTKGAYWRL 506
    480 VTEGGIYVFTLFEDYVAAAGSFLFVFECCIAVAMVYGDIREYDITDIEMLGRGLYKWL 539
QY 507 CMGVIYTPAINTVFYAL-----LASNNLV-FGDNVYVPTAGVSGYILMLFGMTFVP 558
    540 CMKRVSPILTLFELFISIVOGKPLTYNNMIKEADEYVYPNMANALGMLLALSSMLCVP 599
QY 559 IGIGFSLKYRTGTSETIKRA 580
    600 LYIYKILSTEGDSLRLERLQKA 621
Db
```

#### RESULT 4

```
; Sequence 24, Application US/10199485
; Publication No. US2003007626A1
; GENERAL INFORMATION:
; APPLICANT: Glucksman, Maria A.
; TITLE OF INVENTION: 20685, 579, 17114, 23821, 33894, and
; FILE REFERENCE: 32613, No. US2003007626A1el Human Transporters
; CURRENT APPLICATION NUMBER: US/10/199,485
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 09/795,693
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus sequence
US-10-199-485-24
```

Query Match 30.6%; Score 1025.5; DB 15; Length 624;

Best Local Similarity 37.0%; Pred. No. 1.7e-84;

Matches 234; Conservative 94; Mismatches 201; Indels 93; Gaps 15;

```
QY 51 RAVWSNNIEFLMCIATSVGLGNVWRPFPIAYONGGAFVLYVILLGKPYVYECV 110
    1 RETWSGKLDPLVSVGFVAVGLGNVWRPFPIAYONGGAFVLYVILLGKPYVYECV 60
QY 111 LCOFSSRNSVKW-----SISPAKGTGYAQAAGCGYILSYVYVIGCLCYLAMS 161
    61 LQYTRREGSITVWRKKILDKGKICPLFGIGIYASIVIAFYIGIYVNIAMLYLFSS 120
QY 162 FOATLPAICOPEME--NCV-----PSDPTLAA-SYNNITN---GTSSAQLYFLTV 207
    121 FTTELPMATCNNSMTNPNCEVEREAEINSTNSLAALSKNLTDTLERTSPVEEWEKGV 180
QY 208 LQ--OSDGLIEGGLAPIWYVLCFLIAMIWFGVAVGKVS-SGKAAYFLALFPYVMIT 264
    181 LKLSSESGIE-DIGELRWELTCLLAMIWYVFCIMKVGKSGSVVYFTAFPPYVILV 239
QY 265 LEFTTILPGATDGLFEVTPQMAKLELGVWYSANTOVFSLVCTGPTIMESSYNGFR 324
    240 LLIRGVTLPGADGKIFLTLPDSKLLDPQWIDAAQTIFSLGIGFGLIALASYNKFH 299
QY 325 HNIYRDAMIWTTLDFTSFLSGCTIFGLIGNLA----- 357
    300 NNCYRDALIVSFINSITSLAGFVIFSLIGFMANTVOEGVPENEKILLSVLSRDLPH 359
QY 358 -----YEINSEVGD---VVG-----AGTSLAFISYPAIAKTPQPO 391
    360 VNLSTADYDVYDISEVAESEFVLGACLEDELKVOAGPLAFIAPVAVMMLPSP 419
QY 392 LESVLEFLMMSVILGIGSSVALLSTFNLTAMDAFP-----RVPYVMSANTSCGGLLGVY 447
    420 FMAVLEFLMLTLTGIDSGOGVEGIIITLALVDEFPILLRKVRRELFILVCAISFLGLFEM 479
QY 448 CTGPGQYVILEVDHYGTF-FLVFCALSELGAVMIYGLIENICLIDIEFMGKTKGAYWRL 506
    480 VTEGGIYVFTLFEDYVAAAGSFLFVFECCIAVAMVYGDIREYDITDIEMLGRGLYKWL 539
QY 507 CMGVIYTPAINTVFYAL-----LASNNLV-FGDNVYVPTAGVSGYILMLFGMTFVP 558
    540 CMKRVSPILTLFELFISIVOGKPLTYNNMIKEADEYVYPNMANALGMLLALSSMLCVP 599
QY 559 IGIGFSLKYRTGTSETIKRA 580
    600 LYIYKILSTEGDSLRLERLQKA 621
Db
```

#### RESULT 5

```
; Sequence 10, Application US/09843598
; Patent No. US2002010944A1
; GENERAL INFORMATION:
; APPLICANT: Rangachari, H. Robert
; TITLE OF INVENTION: CESTERT GENES, PROTEINS, AND MODULATORY
; FILE REFERENCE: 01997/525002
; CURRENT APPLICATION NUMBER: US/09/843,598
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/200,549
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-843-598-10
```

Query Match

Best Local Similarity 29.5%; Score 991; DB 10; Length 630;

Best Local Similarity 37.0%; Pred. No. 2.3e-81;

Matches 226: Conservative 97: Mismatches 242: Indels 46: Gaps 15:

QY 1 MNDGQVNGFESSEPKMKPRSSDIPPPANNKAALNDIDTDLAE--PPERWWSNNT 58  
 Db 38 VESQOISNGY-SAVPSPGAGDDTRHSIPAT-----TTLVLELHOGHERETGKKV 86  
 QY 59 EFLMSCATSVGICGNWRPFPIAKOEGGAFILPVYVILLVCKPYYLECVIGOPSSRN 118  
 Db 87 DFLSLVGVAVDIGNWRPFPIKQNGGAFILPVYVILLVCKPYYLECVIGOPSSRN 146  
 QY 119 SVKW-SISPMKGTGAQAAGCGYILSYVVICGLLYLAMSFOATLPMACOPWE- 176  
 Db 147 CISTMRKICPIFGIGVIAICITAEFLASYNTIMAMLYLISSTFQDLPTWCKSNMT 206  
 QY 177 -NCVPSDPTLAASVNNIT---NGISKQLYFLTVLO--OSDIEGGLAPITVLCF 230  
 Db 207 GNC-----TNYFSDNTTWTLSHSPAEERYTRVLOIHRSGIQ-DLGISIQALCIM 260  
 QY 231 IAMLVFGVAVARGKSSGKAAYFLAFPPYVMTLFTTLLPGATDGLFEVTPQAKL 290  
 Db 261 LITFVITFSTIKKVKITSGKVVVATPFPYIILSVLVRGATLPGAMGVLEFLKPNQKL 320  
 QY 291 LELGVMSAVTQVFFSLTVCTGPIIMFSSYNGFRHNTYRDAMVITLDTFSLSGCTIF 350  
 Db 321 LETGVWIDAAAOIFFSLGPGCVLILAPASYNKFNNOYODALVTSVNCMTSEFSGVIF 380  
 QY 351 GILGNLAVELNSEVDVVGAGTSLAFISYPDATIAKTFQQLFVLEFLMMSVIGISSV 410  
 Db 381 TVLGYMELMRNEDSEVAKDAGHSLELITTAELANNPASTFPATFIFMLITGLDSTF 440  
 QY 411 ALLSTFNTLANDAPRV-----PTVYMSAMTSCGFLGLVYCTPGQYILELVHGG 464  
 Db 441 AGLEGVITAVLDEPHWAKRRRRVLAIVITCFGSLVLTFF--GGAYVVKLLEYAT 497  
 QY 465 TFLVLCFASISLAVFWIYGLNCLDIEFMLGKKTGAYWRLMGVITPAIMTVFYAL 524  
 Db 498 GPVALTVALLEAVASWYEGITFCGRVDKEMLGSPGFWRICVVAISPLFLITISFL 557  
 QY 525 LASNNL-VFGDNVYPTAGVSYLMFLGMTFVPIGIFSLYKRTGTSETIKKAFHS 583  
 Db 558 MSPQLRLF--QYNPWSIILGKIGTSSFCIPYIAYRLI--ITPOTFERIKK---- 610  
 QY 584 KPSMGPRSPRE 594  
 Db 611 --STPPTPE 619

RESULT 6  
 US-09-843-598-11  
 ; Sequence 11, Application US/09843598  
 ; Patent No. US20020010944A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Horvitz, H. Robert  
 ; APPLICANT: Ranganathan, Rajesh  
 ; TITLE OF INVENTION: CSEBET GENES, PROTEINS, AND MODULATORY  
 ; FILE OF INVENTION: COMPOUNDS  
 ; FILE REFERENCE: 01997/525002  
 ; CURRENT APPLICATION NUMBER: US/09/843,598  
 ; CURRENT FILING DATE: 2001-04-26  
 ; PRIOR APPLICATION NUMBER: US 60/240,549  
 ; PRIOR FILING DATE: 2000-04-26  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 11  
 ; LENGTH: 622  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 ; US-09-843-598-11

Query Match 29.0%; Score 971; DB 10; Length 622;  
 Best Local Similarity 37.5%; Pred No. 1,5e-79;  
 Matches 208; Conservative 97; Mismatches 231; Indels 18; Gaps 8;

QY 51 RMWWSNIEFLMSCATSVGLGNWRPFPIAIONGGAFLVPYVILLVCKPYYLECV 110  
 Db 74 RETMGCKAEFLAVIGFANDLGNWRPFPIKQNGGAFILPVYVILLVCKPYYLECV 133  
 QY 111 LCOFSSRSNVKWS--ISPMKGTGAQAAGCGYILSYVVICGLLYLAMSFOATLPM 169  
 Db 134 LCOFHHCGCLSTMRKICAPLAKGVYALCLIDIMQMYNTLIGMAYVYLFASFISKLPT 193  
 QY 170 ICQPEW--NCVPSDPTLAASVNNITNGTSSAQLFELTVLOOSDIEGGL--GAP1WY 224  
 Db 194 SCQNPNTENCQ-----VTSENFTELATSPAKFEREYVLESTYG--NGIDMGPVAPT 246  
 QY 225 LVLCLEIAMLVFGVAVARGKSSGKAAYFLAFPPYVMTLFTTLLPGATDGLFEVYT 284  
 Db 247 LALCVGVVVLVYFSLMKGVRSAGKVVVATLAPYVILLVVRGSLPGADEGKIYVLT 306  
 QY 285 PQMAKLELGWKSANTQVFFSLTVCTGPIIMFSSYNGFRHNTYRDAMVITLDTFSL 344  
 Db 307 PEWHKTKNSKWIDAAOIFFSLGPGFTLLASYNKFNNOYODALVTSVNCMTSEFSL 366  
 QY 345 SCCTIGILGNLAVELNSEVDVVGAGTSLAFISYPDATIAKTFQQLFVLEFLMMSV 404  
 Db 367 AGVIFSVGLMAYVYKTSI-DKVGLEGVLFVYVPEALAMSGSVFMSIFFLMLTL 425  
 QY 405 GIGSVALLSTENTLANDAPRV---PTVYMSAMTSCGFLGLVYCTPGQYILELVH 461  
 Db 426 GLDSTGEGEAMITLACDEYPRVIGRRRELTVLLAFELFALPTMTGGVAVLVNFW 485  
 QY 462 YGTFVYLCFASISLAVFWIYGLNCLDIEFMLGKKTGAYWRLMGVITPAIMTVFY 521  
 Db 486 YGGGLALFVYVPEAGVFWIYGLNCLDIEFMLGKKTGAYWRLMGVITPAIMTVFY 545  
 QY 522 YALLASNNLVFGDNVYPTAGVSYLMFLGMTFVPIGIFSLYKRTGTSETIKKAF 581  
 Db 546 FSIYKGYEML-GEYIYPPDWSYGVAVYCVSLCIPMYIYKFFPASKGCGQRLOESF 604  
 QY 582 HSKPMGPRSPRE 595  
 Db 605 QPEDNGSVVPGQ 618

RESULT 7  
 US-09-815-923-4  
 ; Sequence 4, Application US/09815923  
 ; Publication No. US20020197644A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gill, Sarjeet S.  
 ; APPLICANT: Ross, Linda S.  
 ; TITLE OF INVENTION: The Regents of the University of California  
 ; FILE OF INVENTION: Target Sites for Insecticides  
 ; FILE REFERENCE: 023070-093800US  
 ; CURRENT APPLICATION NUMBER: US/09/815,923  
 ; CURRENT FILING DATE: 2001-03-23  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 4  
 ; LENGTH: 587  
 ; TYPE: PRT  
 ; ORGANISM: Manduca sexta  
 ; FEATURE:  
 ; OTHER INFORMATION: serotonin transporter  
 ; US-09-815-923-4

Query Match 28.4%; Score 953.5; DB 11; Length 587;  
 Best Local Similarity 36.3%; Pred. No. 5.4e-78;  
 Matches 205; Conservative 106; Mismatches 205; Indels 49; Gaps 12;

QY 18 EPRKSSQISLPPANNKAALNDIDTDLAEPPERWWSNIEFLMSCATSVGLGNWR 77  
 Db 23 QKRSVYSLTPAR-----QREYAKKAEFLAVIGFANDLGNWR 64  
 QY 78 PFIAVONCGAFILPVYVILLVCKPYYLECVIGOPSSRSNVKWS--ISPMKGTGAQ 136



```

Db      65 PYCYONGGAGFLIPYCVMLFEGGLPEFFELALGOYHRCCLITMKRICALKGVYAI 124
Oy      137 AAGGYILSYVYVIGCLLYL---AMSFQATLPAIQCPMEN--CVPSPPTLAASVNN 191
Db      125 CMIDIMKMYNTIIGMAVYLLASINSVLPSTSCDNEMNPCLTP-----VTSPT 179
Oy      192 ITNGTSSAQLYELFTVLOO--SDGIEGLGAPIMYVLCLEFLAMLVGVYARGVSSGK 249
Db      180 NPNSTPAKEFERVLEQHKNSGLD--DMGPDKPSLACVGVFLVYFSLMKGVRSAGK 238
Oy      250 AAYFLALFPYVMTLFTTITLPGATGILFEVYTPQAKLLEGVWTSANTQVFFSLTV 309
Db      239 VVMTALAPYVYVLLILARGVTLPGATEGIRYVLPENHKLKONSQVWIDASQIFESLGP 298
Oy      310 CTGPFMSYNGFPHNITRDAMITVTLDTFTSLGCTIFGILGNLAYELNSEGVYV 369
Db      299 GFGILLALISTNKRNNKCYRDLITSSINCLTSFLAGVITFSLGYMAHVOKST--EEVG 357
Oy      370 AGGTSLAFISYPDAIAKTFQDPSVLFELMMSVIGSSVALLSTFNTLMDAPRY-- 427
Db      358 LEGGLVETIYVPEAIATMTGSVFMAIIEFLMLITGLDSTFGLEAVTALODEYPRVIG 417
Oy      428 --PVTYMSAMTSCGFLGLVYC---TPGGQYILELVHYGTEFLVFCALSELAVF 480
Db      418 RHREFAVVL-----LFIYICAIPTTYGGVYLVLDLLANYGPALILFVFAEAGVC 471
Oy      481 MIYGLENCLEIEFLMGKKTGAYMRLCWGITPAIMTVFFYALLASNNLVFGDYVPT 540
Db      472 WVIYGVDRSEVRYMLTGPWFEMTQMSYISPVFLVLEFVSVAHEML--GGEYTPY 530
Oy      541 AGVYSGYLMFLGMEFVPIGIFSL 565
Db      531 WSIYGVWMTGTIVSCIPLYITIKL 555

```

## RESULT 8

```

US-09-919-039-378
; Sequence 378, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919, 039
; PRIOR FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; NUMBER OF SEQ ID NOS: 2000-07-28
; SOFTWARE: PERL Program
; SEQ ID NO 378
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID NO. US20030108871A1 5834958CD1
US-09-919-039-378

```

```

Query Match      27.8%; Score 932.5; DB 12; Length 614;
Best Local Similarity 33.8%; Pred. No. 4.6e-76;
Matches 208; Conservative 116; Mismatches 241; Indels 51; Gaps 14;
Oy      17 MEPRSSQISLPPA-----NNKKAALNDIDDTLEAEPPEPRMWSNNIEFLMSCIATSYGL 71
Db      1 MDGKVAAYVEYGPRAVSWPBEGERKLDQDEDOVK-----DRGQWTKAMEFVLSVAGEITGL 56
Oy      72 GNVWRPFPIAYONGGAGFLVPYIVYVLLVGRPVYVLECVLQFSSRNSVKYV--SISPAWK 130
Db      57 GNVWRPFPIAYONGGAGFLVPYIVYVLLVGRPVYVLECVLQFSSRNSVKYV--SISPAWK 116
Oy      131 GTGAAAGCGYIISYVYVIGCLCYLYLAMSFOATLPAIQCPEN--ENCVP--SPTPLAA 187

```

```

Db      117 GIGLASVIESYLANVYIIIIILAMALFYLSFSTSELPTTCNNFNTEHCTDFLNHSAG 176
Oy      188 SVNNITNGTSSAQLYELFTVLOOQSDGIEGAGAPIMYVLCLEFLAMLVGVYARGVSS 247
Db      177 TVTPENTSPYMERBERVRLGITGSH--DGLSRLWELALCILLAMVLCYFCIMGVKST 235
Oy      248 GKAAFLALFPYVMTLFTTITLPGATGILFEVYTPQAKLLEGVWTSANTQVFFSL 307
Db      236 GKVVYFTNFPYMLVILLINGVTLPGAYQGIYLYLKDLFRLKQPYQWMAAGRIFFSF 295
Oy      308 TVYCGPITMESSYNGFRNITRDAMITVTLDTFTSFLSGCTIFGILGNLAYELNSEGV 367
Db      296 AICQGLTALSYNNHNNKCYKDCIALCFELNATSFEVAGVFFSILGFMQEOGVPIE- 354
Oy      368 VGAGTSLAFISYPDAIAKTFQDPSVLFELMMSVIGSSVALLSTFNTLMDAPRY-- 426
Db      355 VAESGPGALFIAFPAAVYVMPISQMLSCLEFIMLIFLIDSOFCVBECLVYASIDMFPQ 414
Oy      427 -----PVTYMSAMTSCGFLGLVYCCTPGQYILELVHYGCT--FLVFCALSE 475
Db      415 LRKSGRRRLITLIT---AVMC---YLIGLEVTBEGMYIFQLEFDYASSGICLFLSLFE 468
Oy      476 LAGVYIYGLNCLDIEFLMGKKTGAYWRLCWGITPAIMTVFFYALLASNNLVFGDN 535
Db      469 VVCISWYVYGAURFYDNIEDMIGRWPVLKISMLTEFLGLCLAFELFSLKYYTPLKYNV 528
Oy      536 YVYPTAGVYSGYLMFLGMEFVPIGIFSLKRYRGTSEYTKAF---HSKP----- 585
Db      529 YVYPTAGVYSGYLMFLGMEFVPIGIFSLKRYRGTSEYTKAF---HSKP----- 587
Oy      586 -----SWGPRSPRE 594
Db      588 LDGSAGRNFGPSPTRE 603

```

## RESULT 9

```

US-09-815-923-14
; Sequence 14, Application US/09815923
; Publication No. US2002019764A1
; GENERAL INFORMATION:
; APPLICANT: Gill, Sarjeet S.
; APPLICANT: Ross, Linda S.
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Use of Insect Cell Membrane Transporters as No. US2002019764A
; FILE REFERENCE: 023070-093800US
; CURRENT APPLICATION NUMBER: US/09/815, 923
; NUMBER OF SEQ ID NOS: 2001-03-23
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Manduca sexta
; FEATURE:
; OTHER INFORMATION: GABA transporter
US-09-815-923-14

```

```

Query Match      27.7%; Score 929.5; DB 11; Length 597;
Best Local Similarity 34.0%; Pred. No. 8.4e-76;
Matches 202; Conservative 121; Mismatches 228; Indels 43; Gaps 14;
Oy      17 MEPRSSQISLPPA-----QISLPPANNKKAALNDIDDTLEAEPPEPRMWSNNIEFLMSCIATSYGL 71
Db      1 METKNDSSDIELSAOGSGKRP-----DAVAKSNLEBRGSMASKDPIILISVIGLAIGL 55
Oy      72 GNVWRPFPIAYONGGAGFLVPYIVYVLLVGRPVYVLECVLQFSSRNSVKYV--SISPAWK 131
Db      56 GNVWRPFPIAYONGGAGFLVPYIVYVLLVGRPVYVLECVLQFSSRNSVKYV--SISPAWK 115
Oy      132 TGYAAAGCGYIISYVYVIGCLCYLYLAMSFOATLPAIQCPEN--ENCVP--SPTPLAA 185
Db      116 TGYAAAGCGYIISYVYVIGCLCYLYLAMSFOATLPAIQCPEN--ENCVP--SPTPLAA 175

```

```

0Y 186 AASVNNIT-----NG--TSSAOI- ----YFRTVYQOOSDGGGAGAPFWYVLVLEFJA 232
Db 176 MSSLODMSEFCILNGRNNSKAVLSI PRKEFWERRALOISSGID-HIGNIRWELAGTILLIV 234
0Y 233 WLMVEGVVARGVXSSGKAAVFLTLI PRVVMITLFTITIIILPGATDIDLEFFVTPOMAKLE 292
Db 235 WVLGFCYCIWKRGVWRMTGKVYFTLLI PYFLITVLLIRIGITLLPGAMEGIKFVPMNSKLE 294
0Y 233 LGVWVSATVQVVEFSLTVCTGPIIMI SSYNGFRNITRDMAYITTLDTFFSLSGCIIFGI 352
Db 295 SEWMDAATQOIFESYGLGIGTLTALI GSYNKRFFNNYKDALIYCVSNSSSTMEGAFYISV 354
0Y 353 LGNLAYELNSEEVDVYAGAGTSAIISIYEDATAKTRQPOLESVLEFLMSVLAGISSVAL 412
Db 355 VGFMAHEDQRPAAE-VAASGPGIALI LAYASAVOLPGRAPLMSCLFEFFMLLILDSOFCT 413
0Y 413 LSTFNTLMDADPRV----PTVIMMI AMTSCGFEGLLVCTPGQYIILELVHRYG-GTEL 467
Db 414 MEGFTTAVIDEWPKILLRRKEIITAIITCIISYVLGLOCSISEGGMVFQILDSYVANGFC 472
0Y 468 VLEFCALISELAGFWIYIGLENLCID EFMIGKKTGAWRLCWCYITPALMTTVFFYALLAS 527
Db 473 LLELFFECVSIISMAEGVNRFYIG KEMIGYPTIMKRCWCWGFETPAICISYFIPLNLOW 532
0Y 528 NNLVGDNNVVYPTAGVSGYLMFLF IGMEVPIPIGIGLSLKYRT -GTSEFTIK 579
Db 533 TPIKI-MMEIYMWASHAGCWFTALLI SMLCIP--GMIITLMRVTPGOWDEKFEK 582

```

QY 348 TITGILGNLAYELNSEVDVAGAGTSLAFISYDPAKTFQPOLSVLEFLMMSVLGTC 407  
Db 303 VESVSLGEFAERFNVPVSOVATA--GRLAFVTTPATVTPMAPNLMAITFEVLFELGID 361  
QY 408 SVAALSTFWTLMDAPRVPT--VYNSAMTSCGGLGLVYCTPGQYLTLELVHYGCT 465  
Db 362 TMEVTVLEATLGLDEPRFKSRKRRIAFITCVLFSFSLICTEGGLAHYGLDISHVAI 421  
QY 466 FLV-LFCALSELGAVFMYGLNCLDIEFMLGKKGAVWRKLCMGVITPAIMTVFFYAL 524  
Db 422 LCVPLVCALELIAAV---TYENFSFDVLEMTGRPLRIMVLMRYVILLITLVITL 478  
QY 525 LÄNNLVFGDNVYTPAGVSGYLMFLGMEVPIGFSLYKRTGFSEITKKAHSHK 584  
Db 479 LEVSSL-----AGWFTLVSVICPIYAAVLLR-AGSILLERIRASCRES 523  
QY 585 PSWGRSPRRRRRMMQFKAE 604  
Db 524 NDMGPESEPEKRRMELKKQ 543

RESULT 12  
US-09-843-598-5  
; Sequence 5, Application US/09843598  
; Patent No. US20020010944A1  
; GENERAL INFORMATION:  
; APPLICANT: Horvitz, H. Robert  
; APPLICANT: Ranganathan, Rajesh  
; TITLE OF INVENTION: CESERT GENES, PROTEINS, AND MODULATORY  
; FILE REFERENCE: 01997/525002  
; CURRENT APPLICATION NUMBER: US/09/843,598  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: US 60/200,549  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 671  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-09-843-598-5

Query Match 25.4%: Score 852.5; DB 10; Length 671;  
Best Local Similarity 32.6%: Pred. No. 9.7e-69;  
Matches 193; Conservative 114; Mismatches 256; Indels 29; Gaps 12;

QY 22 SSQISLPPANNKKAALNDIDTD-----LEAEPPERWVMSNNIEFLMSCIATSVGIGNV 74  
Db 64 STSHSIDNPNPIALGGLPTKEGRVALRRSSWVRDKATKMEFLAAYGVAVDLGNT 123  
QY 75 WRFPIAYONGGAFVLYVYVLLVKGPRVYLLVYLLVYLLVYLLVYLLVYLLVYLLVYLLV 133  
Db 124 WRFPSVCKYKHGGAFLLPYFTLMIGLPMFYMELVYGQFHRSGCVSIRKVCPLFRIG 183  
QY 134 YQAAGCGYLLSYVYVIGCLLYLAMS----FOATLPMALCOPEMENCVPSPDTLAASY 189  
Db 184 YGICICTFIAIFNMAIAQAVFAIVSLSKIMSEVPWASCGPWNTPROCDLNVITIS 243  
QY 190 NNITNGSSAQLYFLRVL--QOSDGEGLGAPVLYVLCFLTAMLVGAVARVAKSS 247  
Db 244 RNGPLPTPEEYLLVYLVKLEVOKSTGFD-DLGGVKTSMACVLLAVFIMVYFALMKQPOS 302  
QY 248 GKAAVFLALFPYVVMITLFTTILPGADGILFEFVPMOKKLELGWVSAVYQVFFSL 307  
Db 303 GKIVWVATATAYIILISILIRGLLPKAKGLIYVTPDEKLDPAVWSAAATQIFSL 362  
QY 308 TVCTGPIIMSSYNGFRHNIRYRDAMVITLDTFTSLSGCTIFGILGNLAYELNSEVDV 367  
Db 363 GPGGVLALSSYNDNNNCYRDVAVTIINCATSFSCGVFSTGLYMLTNKRPINEV 422  
QY 368 VGAGGTSLAFISYDPAKTFQPOLSVLEFLMMSVLGSSVAALSTFWTLMD--AFP 425

Db 423 VGEHDSALFIYYPOLATMDYSCWSEFIFVMLITGLDSTFAGIEAFITGFCODESNEL 482  
QY 426 RVPYVMSAMTSCGGLGLVYCTPGQYLTLELVHYGCTFLVLEALSELGAVFYIYGL 485  
Db 483 SKNRKWFVLTITIIYFSLFPAISYGGQVLPFLEDEYGSVLFTVTCENIACVCMFEGV 542  
QY 486 ENCLDIEFMLGKKGAVWRKLCMGVITPAIMTVFFYALLÄN--NLVFGDNVYTPAGY 544  
Db 543 DQFSKDIRAMLEFGYIYRWCV--TCSPVFISVIFIMRYVNSFPIOMASITFEPMWSEVI 601  
QY 545 SGYLMFLGMEVPIGIGFSLKYRTGFSEITKKAHSHKSWGRSPRRR 596  
Db 602 LCMFRLILSVLAIPV---FALYILISGT--GLYERFR---WA-ITPOQR 643

RESULT 13  
US-09-843-598-7  
; Sequence 7, Application US/09843598  
; Patent No. US20020010944A1  
; GENERAL INFORMATION:  
; APPLICANT: Horvitz, H. Robert  
; APPLICANT: Ranganathan, Rajesh  
; TITLE OF INVENTION: CESERT GENES, PROTEINS, AND MODULATORY  
; FILE REFERENCE: 01997/525002  
; CURRENT APPLICATION NUMBER: US/09/843,598  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: US 60/200,549  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 671  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-09-843-598-7

Query Match 25.2%: Score 844.5; DB 10; Length 671;  
Best Local Similarity 32.4%: Pred. No. 5.2e-68;  
Matches 192; Conservative 114; Mismatches 257; Indels 29; Gaps 12;

QY 22 SSQISLPPANNKKAALNDIDTD-----LEAEPPERWVMSNNIEFLMSCIATSVGIGNV 74  
Db 64 STSHSIDNPNPIALGGLPTKEGRVALRRSSWVRDKATKMEFLAAYGVAVDLGNT 123  
QY 75 WRFPIAYONGGAFVLYVYVLLVKGPRVYLLVYLLVYLLVYLLVYLLVYLLVYLLVYLLV 133  
Db 124 WRFPSVCKYKHGGAFLLPYFTLMIGLPMFYMELVYGQFHRSGCVSIRKVCPLFRIG 183  
QY 134 YQAAGCGYLLSYVYVIGCLLYLAMS----FOATLPMALCOPEMENCVPSPDTLAASY 189  
Db 184 YGICICTFIAIFNMAIAQAVFAIVSLSKIMSEVPWASCGPWNTPROCDLNVITIS 243  
QY 190 NNITNGSSAQLYFLRVL--QOSDGEGLGAPVLYVLCFLTAMLVGAVARVAKSS 247  
Db 244 RNTGPTLPTPEEYLLVYLVKLEVOKSTGFD-DLGGVKTSMACVLLAVFIMVYFALMKQPOS 302  
QY 248 GKAAVFLALFPYVVMITLFTTILPGADGILFEFVPMOKKLELGWVSAVYQVFFSL 307  
Db 303 GKIVWVATATAYIILISILIRGLLPKAKGLIYVTPDEKLDPAVWSAAATQIFSL 362  
QY 308 TVCTGPIIMSSYNGFRHNIRYRDAMVITLDTFTSLSGCTIFGILGNLAYELNSEVDV 367  
Db 363 GPGGVLALSSYNDNNNCYRDVAVTIINCATSFSCGVFSTGLYMLTNKRPINEV 422  
QY 368 VGAGGTSLAFISYDPAKTFQPOLSVLEFLMMSVLGSSVAALSTFWTLMD--AFP 425  
Db 423 VGEHDSALFIYYPOLATMDYSCWSEFIFVMLITGLDSTFAGIEAFITGFCODESNEL 482  
QY 426 RVPYVMSAMTSCGGLGLVYCTPGQYLTLELVHYGCTFLVLEALSELGAVFYIYGL 485  
Db 483 SKNRKWFVLTITIIYFSLFPAISYGGQVLPFLEDEYGSVLFTVTCENIACVCMFEGV 542



Mon Jul 28 08:01:33 2003

Search completed: July 21, 2003, 09:37:48  
job time : 60 secs

us-09-991-458-2.rapb

Page 9

**THIS PAGE BLANK (USPTO)**